



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, C12P 21/02 A61K 37/02		A1	(11) International Publication Number: WO 86/ 00639 (43) International Publication Date: 30 January 1986 (30.01.86)
(21) International Application Number: PCT/EP85/00326		136 Wolf Rock Road, Carlisle, MA 017419 (US).	
(22) International Filing Date: 4 July 1985 (04.07.85)		(74) Common Representative: SANDOZ AG; Lichtstrasse 35, CH-4002 Basel (CH).	
(31) Priority Application Numbers: 628,342 652,447 652,742		(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), HU, IT (European patent), JP, KR, LU (European patent), NL (European patent), NO, SE (European patent), SU, US.	
(32) Priority Dates: 6 July 1984 (06.07.84) 19 September 1984 (19.09.84) 19 September 1984 (19.09.84)		<b>Published</b> <i>With international search report.</i>	
(33) Priority Country: US			
(71) Applicant ( <i>for all designated States except US</i> ): SANDOZ AG [CH/CH]; Lichtstrasse 35, CH-4002 Basel (CH).			
(72) Inventors; and			
(75) Inventors/Applicants ( <i>for US only</i> ) : CLARK, Steven, C. [US/US]; 122 Johnson Road, Winchester, MA 01890 (US). KAUFMAN, Randal, J. [US/US]; 111 Marlboro Street, Boston, MA 02116 (US). WONG, Gordon, G. [CA/US]; 1137 Massachusetts Avenue, Cambridge, MA 02138 (US). WANG, Elizabeth, A. [US/US];			

**(54) Title: LYMPHOKINE PRODUCTION AND PURIFICATION**

### (57) Abstract

A method for preparing and isolating a transformation vector containing CSF/cDNA. The method comprises: preparing RNA from a cell that produces CSF; preparing polyadenylated messenger RNA from said RNA; preparing single stranded cDNA from said messenger RNA; converting the single stranded cDNA to double stranded cDNA; inserting the double stranded cDNA into transformation vectors and transforming bacteria with said vector to form colonies; picking pools of 200 to 500 colonies each and isolating plasmid DNA from each pool; transfecting the plasmid DNA into suitable host cells for expressing CSF protein; culturing the transfected cells and assaying the supernatant for CSF activity; and selecting CSF positive pools and screening the colonies used to make the pool to identify a colony having CSF activity. Also described are a cDNA coding for a protein having CSF activity (i.e. CSF/cDNA), a microorganism or cell line transformed with a recombinant vector containing such CSF/cDNA, and a method for producing CSF protein by expressing said CSF/cDNA by culturing a microorganism or cell line. The invention also provides a method of purifying the CSF proteins and the purified proteins so produced.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GA	Gabon	MR	Mauritania
AU	Australia	GB	United Kingdom	MW	Malawi
BB	Barbados	HU	Hungary	NL	Netherlands
BE	Belgium	IT	Italy	NO	Norway
BG	Bulgaria	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	LI	Liechtenstein	SN	Senegal
CH	Switzerland	LK	Sri Lanka	SU	Soviet Union
CM	Cameroon	LU	Luxembourg	TD	Chad
DE	Germany, Federal Republic of	MC	Monaco	TG	Togo
DK	Denmark	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali		
FR	France				

## LYMPHOKINE PRODUCTION AND PURIFICATION

### Field of the Invention

This invention relates to the production of a protein having the ability to stimulate the growth and differentiation of primate hematopoietic progenitor cells, in particular colony stimulating factor (CSF). The invention in one aspect provides a method for producing CSF protein by recombinant DNA techniques, to vectors containing the gene for expressing said protein, to microorganisms and cell lines transformed with said vectors and to CSF protein thus produced. In a second aspect, the invention provides a method for isolating and purifying CSF protein, from either natural or recombinant sources, and thus purified CSF protein having a degree of purity and level of activity well above any that has been previously reported.

### Background of the Invention

The many different cell types found in blood are all derived from pluripotent hematopoietic stem cells. Stem cells perform two functions: (1) they reproduce themselves, thereby maintaining a stem cell population in the body and (2) they provide progeny cells committed to differentiate into any of the mature blood cell types. The cell which is committed to differentiate along a particular hematopoietic pathway is termed a progenitor cell. Progenitor cells for T-lymphocytes, B

lymphocytes, granulocytes, red blood cells, platelets, and eosinophils, as well as earlier progenitors which can individually give rise to several of the mature cell types, have been studied experimentally both *in vivo* and *in vitro* (Dexter, T.M. 1983 *J. Pathology* 141 415-433). It has been determined *in vitro* that proliferation and/or differentiation of each progenitor cell type depends upon specific "factors" which have been derived from various sources. For example, the later progenitors of red blood cells require a factor called erythropoietin. The factors required for survival, proliferation and differentiation of the myeloid progenitors committed to form mature neutrophilic granulocytes, monocytes and mature macrophages are called colony stimulating factors (CSFs).

CSF activity has been studied extensively in the mouse. Most adult mouse organs produce CSF activity. However, compositions containing CSF activity that have been obtained from various tissues and by various methods appear to differ in their biochemical characteristics. Thus, the structural relationships between the different factors remain unknown. Furthermore, CSF activity appears to act at more than one step of granulocyte and macrophage development, and again it has been uncertain whether a single factor is responsible for all of the observed activities or whether a different factor acts at each step. (Burgess, A. and Metcalf, D. 1980 *Blood* 56 947-957).

Human CSF active has been obtained from placenta, certain fetal tissues, macrophages, and stimulated T cells. A line of T

cells (Mo) that produces one or more potent CSF activities was established from a patient with a T cell variant of hairy cell leukaemia (leukaemic reticuloendotheliosis) (Golde et al 1978 Blood 52 1068-1072).

The ability of CS activity to stimulate granulocyte and macrophage production indicated that pharmaceutical compositions having CSF activity are clinically useful in situations where increased production of these (myeloid) cell types is required. Indeed, several patients with extremely high levels of apparently normal circulating granulocytes have been shown to have tumors which over-produce CSF. In one case, upon surgical removal of the tumor, the granulocyte count rapidly declined towards a normal level, strongly suggesting that CSF may be useful in regulating the numbers of circulating granulocytes. (Hocking, W., Goodman, J., and Golde, D. Blood 61 600 (1983)). In particular, CSF compositions are useful clinically for the treatment of myelo-suppression caused by chemotherapeutical or irradiation treatment of cancer. In addition, CSF compositions are useful in treating severe infections because CSF can increase and/or activate the number of granulocytes and/or monocytes.

There are various different types of known CSF activities, including granulocyte CSF (G-CSF), macrophage-CSF (M-CSF), granulocyte-macrophage CSF (GM-CSF) and multi-CSF. The present invention is particularly concerned with GM-CSF. CSF proteins are known from various animal sources. However, the present invention is particularly concerned with primate CSF, more particularly human CSF and ape CSF.

Biological and biochemical characterization of compositions having CSF activity, and study of these compositions in the clinical setting have been hampered to date by the scarcity and impurity of human and/or other primate CSF compositions. It can be appreciated that it would be desirable to identify the protein or proteins responsible for CSF activity. Furthermore, it would be desirable to have a primate, preferably human source of such CSF that could readily supply these proteins in quantities and purity sufficient for biological and biochemical characterization and for use as therapeutic agents.

Recently developed techniques of molecular cloning make it possible to clone a nucleotide sequence which encodes a protein and to produce that protein in quantity using a suitable host-vector system Maniatis, T. Molecular Cloning - A Laboratory Manual Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 1982). The protein can then be recovered by known separation and purification techniques. Cloning methods which have been used to date can be grouped into three general categories: (1) methods based upon knowledge of the protein structure, for example, its amino acid sequence; (2) methods based upon identification of the protein expressed by the cloned gene using an antibody specific for that protein; and (3) methods based upon identification of an RNA species which can be translated to yield the protein or activity encoded by the gene of interest.

Each of these classes of methods becomes difficult to apply when the protein of interest, such as CSF protein, is available in very low amount. Thus, if it is difficult to obtain an adequate quantity of purified protein, then it is difficult to determine the amino acid sequence or even partial sequences of the protein. Similarly, identification of an expressed protein by antibody binding is preferentially carried out using a high-titer monospecific polyclonal anti-serum. Such an antiserum cannot be obtained in the absence of quantities of the pure protein (antigen). A monoclonal antibody offers an alternative approach, but the required antibody can also be difficult to obtain in the absence of suitable antigen, and such monoclonal antibody may not react with the protein in the form in which the protein is expressed by available recombinant host-vector systems. Finally, translation of an RNA species to yield an identifiable protein or activity requires that the RNA in question be present in the RNA source in sufficient abundance to give a reliable protein or activity signal. The relative abundance of an RNA encoding a particular protein generally parallels the abundance of the protein, so that a rare protein is usually encoded by a rare mRNA.

The Mo cell line has been used both as a starting material for purifying human CSFs and for identifying the corresponding messenger RNAs. However, even with this relatively good source of CSF activity, it has proved to be extremely difficult to isolate enough of the protein for structural studies.

In order to overcome the problems inherent in cloning the nucleotide sequence encoding a rare protein such as CSF by the methods described above, a novel method was developed. This method requires only that the gene product or its activity can be

reliably measured. Suitable methods of CSF assay are described in Example 2 hereinafter. In a second aspect, a purification process has been developed which enables the CSF protein to be isolated and purified from either recombinant or natural sources in a level of purity and activity much higher than was previously possible.

Summary of the recombinant DNA process of the Invention

In its first aspect the present invention overcomes the problems of the prior art and provides a ready source of protein having CSF activity using recombinant DNA technology. In accord with the present invention, a novel cloning technique that requires only an assay for CSF activity is utilized to clone cDNA coding for a protein having CSF activity. Thus, the present invention provides a cDNA coding for a protein having CSF activity (i.e. CSF/cDNA), a microorganism or cell line transformed with a recombinant vector containing such CSF/cDNA, and a method for producing CSF protein by expressing said CSF/cDNA by culturing a microorganism or cell line. Because the CSF protein is produced from a clone in accord with the present invention, we can be sure that it is a protein that has CSF activity. The invention further comprises a method for preparing and isolating a transformation vector containing CSF/cDNA, said method comprising:

preparing RNA from a cell that produces CSF;  
preparing polyadenylated messenger RNA from said RNA;  
preparing single stranded cDNA from said messenger RNA;  
converting the single stranded cDNA to double stranded cDNA;

inserting the double stranded cDNA into transformation vectors and transforming bacteria with said vector to form colonies;

picking pools of 200 to 500 colonies each and isolating plasmid DNA from each pool;

transfected the plasmid DNA into suitable host cells for expressing CSF protein;

culturing the transfected cells and assaying the supernatant for CSF activity; and

selecting CSF positive pools and screening the colonies used to make the pool to identify a colony having CSF activity.

The CSF proteins of this invention are growth and differentiation hormones for the cells of the myeloid system. They are for example indicated for use clinically for the treatment of myelo-suppression especially (sympotomatic) granulocyto-penia following chemotherapeutical or irradiation treatment of cancer.

#### Brief Description of the Drawings

Fig. 1 illustrates DNA sequences that code for a CSF protein in accord with the present invention. The DNA sequence set out in full codes for one variation of human CSF, referred to as CSF-Thr. Another allele codes for an identical product except that Thr at position numbered 100 is replaced by Ile (CSF-Ile). The changes illustrated above for the human sequence are for differences in the DNA sequence coding for gibbon CSF (CSF of the Gibbon ape)(CSF-G). Deduced amino acid sequences are also illustrated.

Fig. 2 is a schematic illustrating the preparation of plasmid pTPL from plasmid pAdd26SVpA(3).

Fig. 3 is a schematic continuing from Fig. 2 and illustrating the preparation of plasmid p91023 from plasmid pTPL.

Fig. 4 is a schematic continuing from Fig. 3 and illustrating plasmid p91023(8).

Fig. 6 is a schematic representation of vector pTALC-185R.

Fig. 7 is a schematic representation of vector AJ-14.

#### Detailed Description of the Process

The following definitions are supplied in order to facilitate the understanding of this case. To the extent that the definitions vary from meaning circulating within the art, the definitions below are to control.

Amplification means the process by which cells produce gene repeats within their chromosomal DNA.

CSF is a biological activity defined by the assays as described herein.

CSF protein is a protein from a primate source that exhibits CSF activity. For purposes of the present invention the term CSF protein includes modified CSF protein, allelic variations of CSF protein, and CSF protein preceded by a MET residue.

Downstream means the direction going towards the 3' end of a nucleotide sequence.

An enhancer is a nucleotide sequence that can potentiate the transcription of a gene independent of the position of the enhancer in relation to the gene or the orientation of the sequence.

A gene is a deoxyribonucleotide sequence coding for a given protein. For the purposes herein, a gene shall not include untranslated flanking regions such as RNA transcription initiation signals, polyadenylation addition sites, promoters or enhancers.

Ligation is the process of forming a phosphodiester bond between the 5' and 3' ends of two DNA strands. This may be

accomplished by several well known enzymatic techniques, including blunt end ligation by T4 ligase.

Orientation refers to the order of nucleotides in a DNA sequence. An inverted orientation of a DNA sequence is one in which the 5' to 3' order of the sequence in relation to another sequence is reversed when compared to a point of reference in the DNA from which the sequence was obtained. Such points of reference can include the direction of transcription of other specified DNA sequences in the source DNA or the origin of replication of replicable vectors containing the sequence.

Transcription means the synthesis of RNA from a DNA template.

Transformation means changing a cell's genotype by the cellular uptake of exogenous DNA. Transformation may be detected in some cases by an alteration in cell phenotype. Transformed cells are called transformants. Pre-transformation cells are referred to as parental cells.

Translation means the synthesis of a polypeptide from messenger RNA.

Colony-stimulating factor activity (CSF) can be derived from a number of cellular sources including conditioned medium from peripheral blood mononuclear cells, lung and placental tissue, and bone marrow, urine from anemic patients, serum, and normal and neoplastic cells of T-lymphocyte and mononuclear phagocyte lineage. One cell line that produces CSF is the Mo cell line deposited with and available from ATCC under the code number CRL8066. The CSF produced by this cell line is known as granulocyte macrophage CSF (or GM-CSF) and it is of course a human CSF. One source of Gibbon CSF is the T-cell line designated UCD MLA-144 and deposited with and available from the ATCC under code number HB 9370 deposited September 29, 1983.

In order to isolate a CSF clone in accord with the present invention, a novel procedure was used that requires only an assay technique for CSF activity. First, a cell that produces CSF activity such as T-lymphocyte cells (or other sources such as set forth above) is identified. The mRNA of the cell is then harvested. Preferably, T-lymphocyte cells are used. In such case the membrane bound mRNA, which contains the mRNA for lymphokines, is separated from free mRNA in the cells. This separation is believed to enrich the collected mRNA 5-10 times for lymphokine sequences and thus reduces the effort involved in identifying the desired CSF clone. Polyadenylated messenger RNA is then prepared by chromatography on oligo dT cellulose.

A cDNA library is prepared from the mRNA using a vector suitable for transfection into a host to express the desired protein having CSF activity. First, strand cDNA is prepared using standard methods using the mRNA prepared above. The RNA/cDNA hybrid is then converted to double-stranded cDNA form. The cDNA can then be inserted into a suitable vector.

The preferred host-vector system for the isolation of a CSF clone is based on expression of the CSF cDNA in a suitable transformation vector. A suitable transformation vector can rely on the transient introduction of DNA into mammalian cells (Mellon, P., V. Parker, Y. Gluzman, T. Maniatis 1981 Cell 27 279-288). In order to isolate the desired CSF transformants, it is not required that all cells of the population stably contain exogenous genes that express the desired CSF product. It is possible to transiently introduce exogenous genes into a subpopulation of cells such that the subpopulation will express the desired product over a period of several days. Because a

100-6415

selectable marker is not required in the transformation vector for the DNA transfection and expression system in accord with the present invention, the exogenous DNA can be lost upon growth of the cells over a 1-2 week period. However, 2-3 days after transfection of suitable mammalian cells, the desired products are found to be synthesized and can be detected.

The host-vector system of choice is based on the development of CV-1 monkey cell lines transformed with a replication-origin-defective SV40 DNA molecule (Gluzman, Y., Cell 23 175-182, 1981). The transformed monkey CV-1 cells containing defective SV40 DNA, designated COS (CV-1, origin defective, SV40), do not contain a complete copy of the SV40 genome, but produce high levels of large T antigen and are permissive for SV40 DNA replication. They also efficiently support the replication of SV40 containing deletions in the early region and of bacterial plasmids which contain the SV40 origin of replication (Myers, R.M. & Tjian, R. 1980 PNAS 77 6491-6495). Thus, this system provides a means of amplifying transfected exogenous DNA via SV40 mediated DNA replication in order to increase the level of mRNA and protein expressed from the exogenous DNA. However, other similar systems are also useful.

Vectors used for CSP expression typically contain various elements such as enhancers, promoters, introns, polyadenylation sites, 3' noncoding regions and translational activators as will be described below.

The vectors herein may include enhancers. Enhancers are

functionally distinct from promoters, but appear to operate in concert with promoters. Their function on the cellular level is not well understood, but their unique characteristic is the ability to activate or potentiate transcription without being position or orientation dependent. Promoters need to be upstream of the gene, while enhancers may be present upstream or 5' from the promoter, within the gene as an intron, or downstream from the gene between the gene and a polyadenylation site or 3' from the polyadenylation site. Inverted promoters are not functional, but inverted enhancers are. Enhancers are *cis*-acting, i.e., they have an effect on promoters only if they are present on the same DNA strand. For a general discussion of enhancers see Khoury et al., *Cell* 33:313-314 (1983).

Preferred enhancers for use with mammalian cells are obtained from animal viruses such as simian virus 40, polyoma virus, bovine papilloma virus, retrovirus or adenovirus. Ideally, the enhancer should be from a virus for which the host cell is permissive, i.e. which normally infects cells of the host type. Viral enhancers may be obtained readily from publicly available viruses. The enhancer regions for several viruses, e.g., Rous sarcoma virus and simian virus 40, are well known. See Luciw et al., *Cell* 33:705-716 (1983). It would be a matter of routine molecular biology to excise these regions on the basis of published restriction maps for the virus in question and, if necessary, modify the sites to enable splicing the enhancer into the vector as desired. For example, see Kaufman et al., *J. Mol.*

Biol., 159:601-621 (1982) and Mol. Cell Biol. 2(11):1304-1319 (1982). Alternatively, the enhancer may be synthesized from sequence data; the sizes of viral enhancers (generally less than about 150 bp) are sufficiently small that this could be accomplished practically.

Another element which should be present in the vector assembly is a polyadenylation splicing (or addition) site. This is a DNA sequence located downstream from the translated regions of a gene, shortly downstream from which in turn transcription stops and adenine ribonucleotides are added to form a polyadenine nucleotide tail at the 3' end of the messenger RNA. Polyadenylation is important in stabilizing the messenger RNA against degradation in the cell, an event that reduces the level of messenger RNA and hence the level of product protein.

Eucaryotic polyadenylation sites are well known. A consensus sequence exists among eucaryotic genes: the hexanucleotide 5'-AAUAAA-3' is found 11-30 nucleotides from the point at which polyadenylation starts. DNA sequences containing polyadenylation sites may be obtained from viruses in accord with published reports. Exemplary polyadenylation sequences can be obtained from mouse beta-globin, and simian virus 40 late or early region genes, but viral polyadenylation sites are preferred. Since these sequences are known, they may be synthesized in vitro and ligated to the vectors in conventional fashion.

The sequence which separates the polyadenylation site from

the translational stop codon is preferably an untranslated DNA sequences such as an unpromoted eucaryotic gene. Since such sequences and genes are not endowed with a promoter they will not be expressed. The sequence should extend for a considerable distance, on the order of up to about 1,000 bases, from the stop codon to the polyadenylation site. This 3' untranslated sequence generally results in an increase in product yields. The vector may terminate from about 30 bp downstream from the consensus polyadenylation sequence, but it is preferable to retain the 3' sequences found downstream from the polyadenylation site in its wild-type environment. These sequences typically extend about from 200 to 600 base pairs downstream from the polyadenylation site.

The presence of introns in the untranslated transcribed portion of the vector may increase product yields. Such introns may be obtained from other sources than either the host cells or the gene sources. For example, a hybrid inton comprising a 5' splice site from the second intron of the adenovirus tripartite leader and a 3' splice site from an immunoglobulin gene inserted downstream from transcription start site in the adenovirus major late promoter results in increased product yield.

In the preferred embodiment of the CSF cloning and expression vector there is a translational activator gene. Translational activators are genes which encode either protein or RNA products which affect translation of a desired mRNA. The best example is the adenovirus virus-associated (VA) gene (VAI)

which is transcribed into a short RNA species that interacts with sequences in the 5' untranslated region of the adenovirus major late mRNAs (Thimmappaya et al., 1982 Cell 3 543). The necessary sequences for translational activation by VA RNA lie within the adenovirus late mRNA tripartite leader. The adenovirus tripartite leader is spliced together from noncontiguous regions of the adenovirus genome and is present on the 5' end of the adenovirus major late transcripts. VA RNA can interact to activate translation of mRNAs which contain the tripartite leader sequence. Thus, the preferred cDNA cloning and expression vector contains the spliced form of the tripartite leader, and the adenovirus VA genes.

These vectors can be synthesized by techniques well known to those skilled in this art. The components of the vectors such as enhancers, promoters, and the like may be obtained from natural sources or synthesized as described above. Basically, if the components are found in DNA available in large quantity, e.g. components such as viral functions, or if they may be synthesized, e.g. polyadenylation sites, then with appropriate use of restriction enzymes large quantities of vector may be obtained by simply culturing the source organism, digesting its DNA with an appropriate endonuclease, separating the DNA fragments, identifying the DNA containing the element of interest and recovering same. Ordinarily, a transformation vector will be assembled in small quantity and then ligated to a suitable autonomously replicating synthesis vector such as a prokaryotic

plasmid or phage. The pBR322 plasmid may be used in most cases. See Kaufman et al., *op. cit.*

The synthesis vectors are used to clone the ligated transformation vectors in conventional fashion, e.g. by transfection of a permissive prokaryotic organism, replication of the synthesis vector to high copy number and recovery of the synthesis vector by cell lysis and separation of the synthesis vector from cell debris.

The vectors containing cDNA prepared from a cell that produces CSF activity are then transfected into *E. coli* and plated out on petri dishes at approximately 2000 colonies per dish. The colonies are lifted off onto a nitrocellulose filter and the filter is transferred to a new plate which is kept as a master. After growing these colonies, replicas are made and aligned with the original by careful marking so that sections of the replica filters can be identified with the corresponding portion of the master plate.

Each replica filter is cut into sections containing a predetermined number of colonies per section, preferably about 200-500 colonies per section. The colonies from each section are scraped into medium such as L-Broth, the bacteria collected by centrifugation and the plasmid DNA separated. The plasmid DNA from each section is transfected into a suitable host for expression of protein. The preferred synthesis vector herein is a mutant of the *E. coli* plasmid pBR322 in which sequences have been deleted that are deleterious to eucaryotic cells. See

Kaufman et al., op. cit. Use of this mutant obviates any need to delete the plasmid residue prior to transfection. After growing the transfected cells, the medium is assayed for CSF activity. A positive assay indicates that a colony containing CSF/cDNA is on a particular section of a filter.

To determine which of the clones on the section of the original master filter contains CSF/cDNA, each clone on the filter section is picked and grown. The cultures are then placed in a matrix. Pools are prepared from each horizontal row and vertical column of the matrix. DNA samples are prepared from each pooled culture and transfected into the host cells for expression. Supernatants from these pools are assayed for CSF activity. One vertical column pool and horizontal row pool should produce CSF activity. The clone common to these pools will contain CSF/cDNA. If the matrix contains more than one positive clone, more than one column and row will be positive. In such case, further screening of a small number of clones may be necessary.

The CSF/cDNA is excised from the clones by restriction enzymes and can be sequenced by known techniques. It can be readily appreciated that the procedure described herein can be used to obtain CSF/cDNA from any source. The complete DNA sequence of a CSF/cDNA in accord with the invention is illustrated in Fig. 1 along with the predicted amino acid sequence of the translated CSF protein product.

The DNA sequence coding for a protein exhibiting CSF

activity in accord with the present invention, such as illustrated in Fig. 1, can be modified by conventional techniques to produce variations in the final CSF protein which still have CSF activity in the assay tests described herein. Thus, for example, one, two, three, four or five amino acids can be replaced by other amino acids. Belgian Patent No. 898,016, which is incorporated herein by reference, describes one such typical technique for replacing cysteine by, e.g., serine.

CSF/cDNA in accord with this invention includes the nature CSF/cDNA gene preceded by an ATG codon and CSF/cDNA coding for allelic variations of CSF protein. One allele is illustrated in Fig. 1. Another allele that we discovered has a thymidine residue at position 365 instead of the cytosine residue illustrated in Fig. 1. The CSF protein of this invention includes the 1-methionine derivative of CSF protein (Met-CSF) and allelic variations of CSF protein. The mature CSF protein illustrated by the sequence in Fig. 1 begins with the sequence Ala-Pro-Ala-Arg... the beginning of which is depicted by an arrow after nucleotide number 59 in Fig. 1. The Met-CSF would begin with the sequence Met-Ala-Pro-Ala-Arg... The allele variation illustrated in Fig. 1 has a Thr at amino acid residue number 100 (beginning at Ala after the arrow) and can be referred to as CSF(Thr). Another variation has an Ile residue at position 100 and can be referred to as CSF (Ile). Purified CSF protein of the present invention exhibits a specific activity of at least  $10^7$  units/mg of protein and preferably at least  $4 \times 10^7$  units/mg when assayed with human bone marrow cells.

Host-vector systems for the expression of CSF may be prokaryotic or eucaryotic, but the complexity of CSF may make the preferred expression system a mammalian one. Expression is easily accomplished by transforming prokaryotic or eucaryotic cells with a suitable CSF vector. The DNA sequence obtained by the above described procedure can be expressed directly in mammalian cells under the control of suitable promoter. Heterologous promoters well-known by those skilled in the art can be used. In order to express CSF in prokaryotic or in yeast cells, the leader sequence (or secretory sequence) must be removed. The position of the codon for the N-terminus of the mature CSF protein is illustrated in Fig. 1. This can be done using standard techniques known by those skilled in the art. Once the desired CSF/cDNA clone is obtained, known and appropriate means are utilized to express the CSF protein, e.g. insertion into an appropriate vector, and transfection of the vector into an appropriate host cell, selection of transformed cells, and culture of these transformants to express CSF activity. Suitable host cells include bacteria, e.g. *E. coli*, yeast, mammalian e.g. CHO, and insect cells. The CSF protein thus produced may have a methionine group at the N-terminus of the protein (herein called Met-CSF). The mature protein produced by prokaryotic and eucaryotic cells will be otherwise identical in amino acid sequence, but the eucaryotic product may be glycosylated to the same or a different extent as in the natural product. Various methods of obtaining CSF protein in accordance with the convention are illustrated in the Examples hereinafter. Other methods or materials, e.g. vectors, will be readily apparent to those skilled in the art on the basis of the Examples and the foregoing description.

CSF protein expressed in appropriate prokaryotic or eucaryotic cells can be recovered by purification and separa-

tion techniques known to those skilled in the art. However, as indicated the present invention also provides a purification process which enables CSF protein from both recombinant and natural sources to be obtained in high purity and activity.

#### Summary of the Purification process of the Invention

The present invention overcomes the problems of the prior art and provides a method for purifying protein having CSF activity. CSF protein in accord with the present invention, has specific activity of at least about  $1 \times 10^7$  units per mg of protein, preferably at least  $2 \times 10^7$  units per mg of protein and more preferably at least about  $4 \times 10^7$  units per mg of protein when assayed the human bone marrow assay.

In accord with the present invention, a method for purifying CSF protein comprises: precipitating the protein with ammonium sulfate at 80% saturation to form a pellet containing the CSF protein; resuspending the pellet in a buffered solution at a pH in the range of about 6 to about 8; applying the buffered solution containing CSF to a chromatographic column, eluting with the buffered solution containing sodium chloride and collecting the fractions having CSF activity; pooling the active fractions, applying them to a C4 reverse phase column and eluting with a 0 to 90% acetonitrile gradient to collect the active fraction.

#### Brief Description of the Drawings relating to the purification process

Fig. 5 illustrates SDS-PAGE analysis of the purified CSF protein.

#### Detailed Description of the purification process of the Invention

The CSF protein to be purified in accordance with the process of the invention can be derived from any of the natural sources described above as starting sources for the recombinant

DNA process, for example the Mo cell line or the UCD MLA-144 Gibbon cell line.

Alternatively, the CSF protein may be produced using the recombinant DNA techniques of the invention.

CSFs from any source can be purified by the process of the present invention. The conditioned medium from any source of CSF protein is preferably concentrated by ultrafiltration to a protein concentration of at least about 0.1 mg protein per ml. The protein is then precipitated by adding ammonium sulfate to 80% of saturation. The resulting pellet is resuspended in an aqueous solution buffered at a pH in the range of about 6 to about 8. Examples of suitable buffers include Tris-HCl, HEPES, sodium citrate, and the like.

The buffered solution is fractionated by column chromatography. Suitable materials for use in the chromatography column are octylsepharose, DEAE-ultrogel, AcA44-ultrogel, AcA-54 ultrogel, and the like. One or more of these materials can be used in sequence to obtain higher purity.

Fractions from each column are collected and assayed for CSF activity. The active fractions are pooled and diluted with trifluoroacetic acid (TFA), heptafluorobutyric acid (HFBA), or the like, and applied to a C4 reverse phase column. The CSF activity is then eluted using a 0-90% acetonitrile gradient in TFA or HFBA, preferably at a concentration of 0.10% or 0.15% (vol/vol) respectively, depending upon which acid was used to apply the pooled fractions to the column.

The fractions having CSF activity are analyzed by SDS polyacrylamide gel electrophoresis (13.5% gel as described by Lammli, U. Nature 227, 680 (1970)). Additional treatments using the above mentioned chromatographic column materials can further purify the CSF protein to homogeneity.

Purified CSF protein fractionated by SDS-PAGE revealed a heterogeneous CSF protein having an apparent molecular

weight in the range of about 15,000 to about 26,000 daltons. This apparent size heterogeneity is due to the extensive glycosylation of the protein and is a common feature of glycoproteins. Fractionation of less purified samples from Mo cell conditioned medium by SDS-PAGE (under non-reducing conditions) and assaying protein eluted from the gel revealed the presence of a second protein having CSF activity having an apparent molecular weight of about 28,000 to 30,000.

CSF activity binds and elutes from octylsepharose, DEAE ultrogel and the C4 reverse phase column. Roughly 60% of the CSF activity binds a Con-A sepharose (40% flow through) and can be eluted with alpha methylmannoside.

Molecular weight analysis of recombinant CSF by gel filtration in low salt revealed that about 30% of the activity eluted with an estimated molecular weight of about 19,000 but 70% of the material behaved as dimers, eluting at a position corresponding to a molecular weight of about 38,000. If 1M NaCL is included in this column, all of the activity elutes in a broad peak at about 19,000 daltons.

The purified CSF is stable for at least 16 hours when incubated at 4°C (pH 7.4) in 4M guanidine hydrochloride; in 10mM EDTA; 10mM 2-mercaptoethanol; and in 30% (v/v) ethanol. The CSF activity also is stable in 0.1% trifluoroacetic acid (TFA) (pH 2.0) and 0.1% TFA plus 25% (v/v) acetonitrile.

As aforesaid, the CSF protein in accord with the present invention is indicated for use in the treatment of myelosuppression such as (symptomatic) granulocytopenia, for example caused by chemotherapeutical or radiation treatment of cancer. In addition, CSF proteins of the invention are indicated for use in the treatment of severe infection. For such use, an indicated dosage of about 200 to 1000 ug per patient is typically indicated. The CSF protein is preferably injected into the patient intravenously in a suitable pharmacological carrier. Examples of such carriers include pharmacological saline and human serum albumin in saline.

In addition, the CSF proteins of the invention have other activities and uses. For instance, it has been shown that murine CSFs activate neutrophils. Thus, it would be expected that the primate CSFs of the present invention will also activate neutrophils. Therefore, physiological functions of CSF may be severalfold. In the bone marrow, this lymphokine can stimulate proliferation and differentiation of effector cells for host defense while, in the periphery, new and existing cells can be activated. In a localized immunological response CSF can retain circulating neutrophils in or away from areas of inflammation. Inappropriate localization and/or activation of neutrophils can be involved in the pathophysiology of a variety of immune-mediated disorders such as rheumataid arthritis.

The invention will be further understood with reference to the following illustrative embodiments, which are purely exemplary, and should not be taken as limitative of the true scope of the present invention, as described in the claims.

In the examples, unless otherwise specified, temperatures are in °C.

Restriction endonucleases are utilized under the conditions and in the manner recommended by their commercial suppliers. Ligation reactions are carried out as described by Maniatis et al., supra at 245-6, the disclosure of which is incorporated herein by reference, using the buffer described at page 246 thereof and using a DNA concentration of 1-100 µg/ml, at a temperature of 23°C for blunt ended DNA and 16°C for "sticky ended" DNA. Electrophoresis is done in 0.5-1.5% Agarose gels containing 90 mM Tris-borate, 10 mM EDTA. All radiolabeled DNA is labeled with <sup>32</sup>P, whatever labeling technique was used.

By "rapid prep" is meant a rapid, small scale production of bacteriophage or plasmid DNA, e.g., as described by Maniatis et al., supra, at p. 365-373. ERSATZBLATT

EXAMPLE AStep 1. Mo Cell Line Cultures

Mo cells (ATCC CRL 8066) were grown routinely in Alpha (6% CO<sub>2</sub>) or Iscove's (10% CO<sub>2</sub>) medium containing 20% Fetal Calf Serum (FCS), 2mM glutamine, 100 U/ml streptomycin and 100 ug/ml penicillin. The cells should be subcultured every 4-5 days. Cells are counted and seeded into Falcon T-175 flasks in 100-150 ml medium at density of 3-4 x 10<sup>5</sup> cells/ml. Cells will double in 20% FCS every 4-7 days. Growth rate is not constant and cells may sometimes appear to stop growing then go through bursts of growth. Mo cells can be grown in serum-free medium. Survival is much better when cells are not washed when transferred from FCS to serum-free medium. Optimal density in Serum-Free medium (SF) is 5 x 10<sup>5</sup> cells/ml. Cells will grow slightly (or at least maintain constant number) for 3 days in serum-free medium, and then should be fed 20% FCS for at least 4 days. This growth schedule (3 days SF, 4 days 20% FCS) can be repeated weekly if SF medium is required, with no apparent harm to the cells for several months.

Step. 2 Assays for CSF Activity

## A. Bone Marrow Assay

Obtain fresh bone marrow. Break apart spicules by drawing through 20, 22, then 25 gauge needle. Dilute 1:1 with sterile phosphate-buffered saline (PBS) (room temperature) and layer over Ficoll-Paque (about 30 ml BM-PBS over 6 ml Ficoll). Spin at 1500

rpm for 40 minutes at room temperature. Remove fat and PBS layer and discard. Pipette off the light density layer. Wash 2x with PBS and count. Plate cells in RPMI (purchased from GIBCO as RPMI 1640) plus 10% HIFCS (heat inactivated FCS) for 3 hours to remove adherent cells.

Plating medium (make fresh):

20% FCS

0.3% agar dissolved in H<sub>2</sub>O cooled to 40°C

2x Iscoves (1:1 v/v with Agar)

1% P/S final concentration of 100U/ml streptomycin,

100 ug/ml penicillin

10<sup>-4</sup>M alpha thioglycerol in 2x iscoves from 10<sup>-2</sup>M

stock

Cool agar to about 40°. Mix with other ingredients.

Cool in H<sub>2</sub>O bath to 37-38° and hold at that

temperature.

After 3 hours, pipette off the non-adherent cells. Spin and count. Add 2 x 10<sup>5</sup> cells/ml of plating medium and keep in controlled temperature water bath at 37-38°. Add samples (e.g., medium from transfected cells; usually 10  $\mu$ l sample) to the first row of wells of a microtiter plate in duplicate. Add 100  $\mu$ l cell suspension to each well. Add additional 50  $\mu$ l of cell suspension to each well in the first row. Mix thoroughly and transfer 50  $\mu$ l of solution from the first row into the next row, etc. and continue 1:3 dilutions across plate. Wrap the plate in parafilm.

Incubate 10-14 days at 10% CO<sub>2</sub>, 37°C in fully humidified atmosphere and score colonies.

To score the colonies, the total number of colonies that grow in each well is counted. In each assay, several wells are plated without including a sample (blank) to obtain a background colony count. The average number of colonies that grew in the blank wells is subtracted from the number of colonies found in each of the wells containing samples. One unit of CSF is the amount that will stimulate the formation of one colony above the background level per 10<sup>5</sup> human bone marrow cells (plated at 10<sup>5</sup> cells per ml) when the CSF concentration is sub-saturating. The sub-saturating concentration is determined by dilution and comparing the number of colonies at various dilutions to find the concentration just below the saturation level.

For this assay, the colonies containing granulocytes, monocytes or both types of cells are counted. The types of cells in the colonies are determined by picking colonies and staining individual cells.

#### B. KG-1 Cell Assay

KG-1 cells (Blood, Vol. 56, No. 3 (1980)) are grown in Iscoves medium + 10% FCS passed 2x per week and seeded for each passage at 2x10<sup>5</sup> cells/ml. The cells are used for assay only between passage 30-35. The assay is the same as for bone marrow as described above, except the KG-1 cells are plated in agar mixture at 4x10<sup>3</sup> cells/ml.

The number of colonies growing in each well is determined and the background count is subtracted as in the Bone Marrow assay described above. One KG-1 CSF unit/ml is that concentration of CSF that will stimulate half of the maximum number (saturation) of KG-1 colonies to grow. The maximum number is obtained by including a saturating level of CSF in several wells.

Step 3. Construction of Vector p91023(B)

The transformation vector was pAdD26SVpA(3) described by (Kaufman et al., Mol. Cell Biol. 2(11):1304-1319 (1982). It has the structure illustrated in Fig. 2. Briefly this plasmid contains a mouse dihydrofolate reductase (DHFR) cDNA gene that is under transcriptional control of the adenovirus 2 (Ad2) major late promoter. A 5' splice site is included in the adenovirus DNA and a 3' splice site, derived from an immunoglobulin gene, is present between the Ad2 major late promoter and the DHFR coding sequence. The SV40 early polyadenylation site is present downstream from the DHFR coding sequence. The prokaryotic-derived section of pAdD26SVpA(3) is from pSV2d (Mellen, P., Parker, V., Gluzman, Y. and Maniatis, T. 1981, Cell 27:279-288) and does not contain the pBR322 sequences known to inhibit replication in mammalian cells (Lusky, M., and Botchan, M. 1981, Nature (London) 291:79-81.

pAdD26SVpA(3) is converted into plasmid pCVSVL2 as illustrated in Fig. 2. pAdD26SVpA(3) is converted into plasmid

pADD26SVpA(3)(d) by deletion of one of the two *Pst*I sites in pADD26SVpA(3). This is accomplished by a partial digestion with *Pst*I (using a deficiency of enzyme activity so that a subpopulation of linearized plasmids can be obtained in which only one *Pst*I site is cleaved), then treatment with Klenow, ligation to recircularize the plasmid, transformation of *E. coli* and screening for deletion of the *Pst*I site located 3' of the SV40 polyadenylation sequence.

The adenovirus tripartite leader and virus associated genes (VA genes) were inserted into pADD26SVpA(3)(d) as illustrated in Fig. 2. First, pADD26SVpA(3)(d) was cleaved with *Pvu*II to make a linear molecule opened within the 3' portion of the first of the three elements comprising the tripartite leaders. Then, pJAH 43 (Zain et al. 1979, Cell 16 851) was digested with *Xba*I, treated with Klenow, digested with *Pvu*II, and the 140 base pair fragment containing the second and part of the third leaders was isolated by electrophoresis on an acrylamido gel (6% in Tbis borate buffer; Maniatis et al. (1982) supra). The 140 bp fragment was then ligated to the *Pvu*II digested pADD26SVpA(3)(d). The ligation product was used to transform *E. coli* to tetracycline resistance and colonies were screened using the Grunstein-Hogness procedure using a <sup>32</sup>P-labelled probe hybridizing to the 140 base pair fragment. DNA was prepared from positively hybridizing colonies to test whether the *Pvu*II site reconstituted was 5' or 3' of the inserted 140 base pair DNA specific to the 2nd and 3rd adenovirus late leaders. In the correct orientation of the *Pvu*II

site is on the 5' side of the 140 base pair insert. This plasmid is designated pTPL in Fig. 2.

The Ava II D fragment of SV40 containing the SV40 enhancer sequence was obtained by digesting SV40 DNA with Ava II, blunting the ends with Klenow fragment of Pol I, ligating Xho 1 linkers to the fragments, digesting with Xho 1 to open the Xho 1 site, and isolating the fourth largest (D) fragment by gel electrophoresis. This fragment was then ligated to Xho 1 cut pTPL yielding the plasmid pCVSL2-TPL. The orientation of the SV40 D Fragment in pCVSL2-TPL was such that the SV40 late promoter is in the same orientation as the adenovirus major late promoter.

To introduce the adenovirus virus associated (VA) genes into the pCVSL2-TPL, first a plasmid pBR322 is constructed that contains the adenovirus type 2 Hind III B fragment. Adenovirus type 2 DNA is digested with Hind III and the B fragment is isolated after gel electrophoresis. This fragment is then inserted into pBR322 which has previously been digested with Hind III. After transformation of *E. coli* to ampicillin resistance, recombinants are screened for insertion of the Hind III B fragment and the inserted orientation is determined by restriction enzyme digestion. pBR322 - Ad Hind III B contains the adenovirus type 2 Hind III B fragment in the orientation depicted in Fig. 3.

As illustrated in Fig. 3, the VA genes are conveniently obtained from plasmid pBR322-Ad Hind III B by digesting with Hpa

I, adding EcoRI linkers and digesting with EcoRI, and recovering the 1.4kb fragment. The fragment having EcoRI sticky ends is then ligated into the EcoRI site of pTEL (which had previously been digested with EcoRI). After transformation of E. coli HB101 and selection for tetracycline resistance, colonies are screened by filter hybridization to a DNA probe specific to the VA genes. DNA is prepared from positively hybridizing clones and characterized by restriction endonuclease digestion. The product plasmid is designated p91023.

The 2 EcoRI sites in p91023 are removed. p91023 is cut to completion with EcoRI, generating two DNA fragments, one about 7Kb and the other about a 1.3 Kb fragment containing the VA genes. The ends of both fragments are filled in using the Klenow fragment of PstI, and then both fragments i.e. 1.3 Kb, 7Kb, are re-ligated together. A plasmid p91023(A) containing the VA genes and similar to p91023 but deleted for the 2 EcoRI sites is identified by Grunstein-Hegness screening with the VA gene fragment, and by conventional restriction site analysis.

Then the single PstI site in p91023(A) is removed and replaced with an EcoRI site. p91023(A) is cut to completion with PstI, and then treated with Klenow fragment of PstI to generate flush ends. EcoRI linkers are ligated to the blunted PstI site of p91023(A). The linear p91023(A), with EcoRI linkers attached to the blunted PstI site is separated from unligated linkers and digested to completion with EcoRI, and then re-ligated. A plasmid p91023(B) is recovered and identified to have a structure similar

to p91023(A), but with an EccR1 site situated at the previous pstI site.

Step 4. Preparation of cDNA Library

No cells were induced for 16-20 hrs. with PHA and PMA to enhance their lymphokine production. Cells were plated at 5 x 10<sup>5</sup> cells/ml in Iscove's medium with 20% FCS, 0.3% (v/v) PHA and 5 ng/ml TPA. The cells were collected by centrifugation. The pelleted cells were resuspended in 20 ml of ice cold hypotonic lysis buffer (RSB buffer: 0.01M Tris-HCl, pH 7.4, 0.01M KCl, 0.0015M MgCl<sub>2</sub>, 1 ug/ml cycloheximide, 50 units/ml RNAsin and 5mM dithiothreitol). The cells were allowed to swell on ice for five minutes then were ruptured mechanically with 10 strokes of a tight fitting dounce glass homogenizer. The homogenate was centrifuged at low speed (2000 RPM in a Beckman J6 centrifuge) to remove nuclei and unlysed cells. The supernatant was held on ice while the nuclear pellet was resuspended in 10 ml of RSB and re-centrifuged at low speed. This second supernatant was pooled with the first and the combined supernatants were centrifuged at low speed to remove residual contamination with nuclei and unlysed cells. The supernatant from this spin was brought to 0.15M KCl by addition of 2M KCl then centrifuged at high speed (25,000 RPM, Beckman SW 28 rotor for 30 minutes) to pellet the membranes. The membrane pellet was carefully washed with cold RSB then resuspended in 12 ml of RSB containing 2 M sucrose and 0.15M KCl. Two discontinuous gradients were prepared in Beckman

SW41 centrifuge tubes by laying 6 ml of the membrane solution in 2 M sucrose over 2 ml of RSB with 2.5 M sucrose and 0.15M KCl. The tubes were filled to the top by overlaying with 2.5 ml of RSB containing 1.3M sucrose and 0.15M KCl. These gradients were spun for 4 hours at 27,000 RPM (Beckman, SW41 rotor) at 4°C. The membrane layer (at the interface between the 2.0M and 1.3M sucrose) was carefully removed from the side using an 18 gauge needle and syringe. The membrane fractions from the two gradients were pooled and diluted with 1 volume of distilled H<sub>2</sub>O then brought to 0.5% Triton X-100 and 0.5% sodium deoxycholate then extracted with an equal volume of phenol. The aqueous layer was re-extracted with a 1:1 mixture of phenol and chloroform and finally an equal volume of chloroform. Finally, the membrane bound RNA was precipitated by addition of NaCl to 0.25M and 2.5 volumes of cold ethanol and incubated overnight at -20°C. The precipitated RNA was collected by centrifugation (4000 RPM for 10 min. in the Beckman J-6 centrifuge) and was resuspended in 1 ml of distilled water. From 2x10<sup>9</sup> cells, approximately 1 mg of RNA was obtained. The messenger RNA (mRNA) was isolated from the total RNA by chromatography on a 0.5 ml oligo dT-cellulose column. Briefly the RNA was heated to 70°C for 5 min., quick chilled on ice, then diluted 5 fold with room temperature binding buffer (0.5M LiCl, 0.01M Tris-HCl, pH 7.4, 0.002 M EDTA, and 0.1% SDS). The RNA in binding buffer was passed over the oligo dT-cellulose column equilibrated with binding buffer at room temperature. The column was washed with 5 ml of binding buffer.

then with 5  $\mu$ l of 0.15M LiCl, 0.01M Tris-HCl pH7.4, 0.002M EDTA, and 0.1% SDS. Finally, mRNA was eluted with 2 ml of 0.01M Tris-HCl pH7.4, 0.002M EDTA, and 0.1% SDS. The mRNA was precipitated by addition of NaCl to 0.25 M and 2.5 volumes of ethanol and incubation overnight at -20°C. The precipitated mRNA was collected by centrifugation (30,000 RPM for 30 minutes in a Beckman SW55 rotor). The tube was carefully drained and the mRNA pellet was resuspended in 50 ml of  $H_2O$ . The resuspended mRNA was brought to 0.25M NaCl then extracted 1 time with a 1:1 mix of phenol and chloroform then 3 times with chloroform. The mRNA was precipitated by the addition of 2.5 volumes of ethanol. The mixture was frozen and thawed several times in a dry ice/ethanol bath then centrifuged 15 min. in an Eppendorf centrifuge. The tube was carefully drained and the mRNA pellet was resuspended in 20  $\mu$ l of distilled  $H_2O$ . The final yield was approximately 30  $\mu$ g of mRNA.

First strand cDNA was prepared using standard methods. Briefly, 10  $\mu$ g of membrane mRNA was diluted into a 100  $\mu$ l cDNA synthesis reaction mixture containing 300 mM Tris pH 9.4, 140 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, 500  $\mu$ M each of dATP, dGTP, dCTP and dTTP, 5  $\mu$ g. of oligo-dT (phosphorylated and average size of 12-18) as primer, 150  $\mu$ Ci of <sup>32</sup>P dCTP (400 Ci/mmol) and 20 units of the ribonuclease inhibitor RNasin. The reaction was initiated by addition of 100 units of reverse transcriptase and incubated for 30 minutes at 42°C. The reaction was stopped by addition of EDTA to 40 mM and the RNA was degraded

by incubation for 20 min. at 65°C in 0.2M NaOH. The base was neutralized by addition of 20 ul 2M Tris, pH 7.4. The reaction mix was then extracted with phenol/chloroform, back extracted with 50 ul 10 mM Tris pH 7.5, 1 mM EDTA (TE) and the aqueous phases were pooled. The first strand cDNA was converted to double stranded cDNA by incubation for 12 hours at 16°C with 40 units of the Klenow fragment of DNA polymerase I in a 100 ul reaction containing 50 mM potassium phosphate, pH 7.4, 2.3 mM DTT, 2-mercaptoethanol, 10mM MgCl<sub>2</sub>, 150 uMolar each of the 4 deoxynucleoside triphosphates and 25 uCi of <sup>32</sup>P dCTP. The reaction was stopped by extraction with phenol/chloroform and the unincorporated triphosphates were removed by passing the aqueous phase over a 1 ml sephadex G-50 column. The excluded fractions were pooled and ethanol precipitated.

The cDNA pellet was washed with cold ethanol then resuspended in 200 ul of 20mM Tris pH 6.0, 1 mM EDTA, 80uMolar S-adenosyl-Methionine, and 300 units of EcoRI methylase for 60 minutes at 37°C. The reaction was stopped by extraction with phenol/chloroform and the methylated cDNA was collected by ethanol precipitation.

The cDNA pellet was rinsed with 700 ethanol then resuspended in 200 ul S1 buffer (Maniatis et al) and incubated with 200 units of S1-nuclease at 30°C for 30 minutes. The reaction was stopped by extraction with phenol/chloroform and the cDNA collected by ethanol precipitation.

The double stranded cDNA was blunted by incubation in 100 ul

of 20mM Tris, pH 7.4, 50mM NaCl, 10mM 2 mercaptoethanol and 500  $\mu$ Molar of all four deoxynucleotide triphosphates with 25 units of Klenow at room temperature for 30 minutes. The reaction was stopped by extraction with phenol/chloroform and the cDNA collected by ethanol precipitation.

The cDNA was ligated in 50 ul of T4 ligase buffer (Maniatis et al) with 500 pMoles of R1 linkers purchased from New England Biolabs (sequence: pCGGAATTCCG) using 2000 units of T4 ligase overnight at 16°C. The reaction was stopped by incubation at 70° for 20 minutes then diluted to 300 ul such that the final salt concentration was 0.1 M NaCl, 10 mM, MgCl<sub>2</sub>, 50 mM Tris-Cl, pH 7.4. The cDNA was then digested for 2 minutes at 37° with 700 units of EcoRI. The reaction was stopped by extraction with phenol/chloroform and the cDNA collected by ethanol precipitation. The pellet was resuspended in 50 ul of TE and passed over a 5ml Cl-4B column. The excluded fractions were pooled and ethanol precipitated. The precipitated cDNA was electrophoresed through a 1% agarose gel in Tris acetate buffer in the presence of 1 ug/ml ethidium bromide. cDNA in the size range 500-4000 base pairs was isolated from the gel using the standard glass powder procedure. The eluted cDNA was extracted with phenol/chloroform, ethanol precipitated and the pellet (after an ethanol rinse) was resuspended in 30ul of TE. The final yield was 100-500 ng of cDNA.

The preparation of the expression vector p91023(B) is described above. The EcoRI digested and phosphatase treated

vector (500ng) was ligated with 100 ng of cDNA in a 100ul reaction (standard T4 ligase reaction) overnight at 16°C. The reaction was stopped by extracting with phenol/chloroform then the ligated cDNA was collected by ethanol precipitation after adding 5 ug of tRNA as carrier.

The ethanol precipitated DNA was rinsed with 70% ethanol then resuspended in 100 ul of TE. This DNA was used in 4 ul aliquots to transform *E. coli* MC1061 (4 ul in a 100 ul transformation). Each of the 25 transformations was spread onto a 150 mm petri dish with 1% agar, L-broth and 10 ug/ml tetracycline (Tet plate) and incubated overnight at 37°. Approximately 2000 colonies grew on each plate, resulting a total of about 50,000 colonies. After reaching approximately 0.5 mm in diameter, the colonies were transferred to nitrocellulose disks (137 mm) by carefully placing a dry filter on the surface of the plate then smoothly peeling off the filter. All of the colonies on the plate transferred to the filter which was then placed (colony side up) on a fresh Tet plate. After allowing the colonies to grow several hours, one replica was prepared from each of the filters by placing a fresh wetted filter exactly over the original filter, pressing them together, peeling them apart then returning each filter to a fresh Tet plate and incubating the plates overnight at 37°. Each replica was carefully marked such that it would be realigned with the original filter.

### Step 5. Plasmid DNA Pool Preparation

Each of the 25 replica filters was carefully sectioned into eighths using a scalpel and noting the orientation of each eighth relative to the original master filter. The colonies were scraped from each section into 10 ml of L-Broth. The bacteria were collected by centrifugation (3000 RPM, 10 min., Beckman J-6. centrifuge) resuspended in 0.6 ml of 25% sucrose, 50 mM Tris-HCl, pH 8.0 and converted to protoplasts by addition of 0.12 ml of 5 mg/ml lysozyme and incubation on ice for 5-10 min. The protoplasts were next incubated at room temperature for 10 min. following the addition of 0.125 ml of 0.5M EDTA then lysed by addition of 0.12 ml of 10% SDS in 50 mM Tris-HCl, pH 8.0. The lysate was mixed gently, incubated at room temperature for 15 min. then protein and chromosomal DNA precipitated by the addition of 0.3 ml of 5M NaCl. After incubation on ice for 15 min., the lysate was centrifuged in an Eppendorf centrifuge for 30 min. in the cold. The supernatant was carefully removed leaving behind the viscous DNA/protein pellet and was diluted by the addition of 2.5 ml H<sub>2</sub>O. The mixture was extracted with 1 ml of phenol, the layers separated by centrifugation (10K for 10 min. in the Sorvall SS-34 rotor) and the aqueous layer removed to a fresh tube. DNA was precipitated by adding 0.5 ml of 5M NaCl and 7.5 ml of cold ethanol and freezing the mixture several times in a dry ice ethanol bath. The precipitate was collected by centrifugation (10K, 15 min. in the Sorvall SS-34), resuspended

in 0.3 ml of 0.3M Sodium acetate and re-precipitated (in an Eppendorf tube) by the addition of 1 ml of ethanol. After 10-15 min. in a dry ice ethanol bath, the precipitated DNA was collected by centrifugation (5 min. in the Eppendorf) and the final pellet was resuspended in 100 ul of sterile TE (10 mM Tris pH8, 1mM EDTA). From a typical preparation, 5-10 ug of plasmid DNA was obtained. Each preparation contained the DNA from 200-500 colonies on the original filter. A total of 200 DNA samples were prepared from the 25 filters.

#### Step 6. Isolating CSF Clone

Each of the DNA samples from Step 5 were separately transfected into M6 COS monkey cells as described below.

The M6 cells are grown routinely in Dulbecco's modified Eagle's Medium (DME available from Gibco) containing 10% heatinactivated fetal calf serum (HIFCS), split twice a week at 1:6 dilution. Twenty-four hours after splitting 1:6 the M6 cells are ready for transfection. Twenty-four hours prior to transfection,  $1.2 \times 10^8$  M6 cells (split 1:6) are seeded into a Cell Factory (available from Nunc) in 1.5 liters of DME + 10% HIFCS. Immediately before transfection, plates are aspirated and washed twice with 7 ml of serum-free (SF) DME. The DNA is dissolved in 0.1 M Tris (pH 7.3) and added to DME medium containing 2mM glutamine, 100 ug/ml streptomycin, 100 U/ml penicillin and 0.25 mg/ml DEAE Dextran totalling 4 ml with the Tris-DNA solution. The 4 ml of medium containing dissolved DNA is added to the plate containing M6 COS cells and incubated for 12 hours.

After incubation, the cells are rinsed once or twice with 7 ml SF DME. Then, 5 ml of DME with 10% HIFCS, 100 U/ml penicillin, 100 ug/ml streptomycin, 2mM glutamine, and 0.1 mM chloroquin was added and the cells were incubated for 2 1/2 hours.

ERSATZBLATT

After 2 1/2 hours, rinse once with SF DME and add 10 ml DME + 10% HIFCS/plate. After 30 hours aspirate media and feed 4 ml/plate DME + 10% HIFCS. Harvest by removing the conditioned medium after 24-26 hours further incubation.

The conditioned medium from each transfection was assayed for CSF activity using the KG-1 assay. For each sample, positive for CSF activity, the clone on the original master filter responsible for the CSF activity had to be identified. For example, for one transfection positive for CSF activity, all of the colonies of the section of the original master filter from where the transfection DNA sample was derived, were picked. Some 320 of these colonies were picked into 3 ml of L-Broth plus 10 ug/ml tetracycline. The cultures were grown overnight. The 320 colonies were placed in an 18 x 18 matrix. Pools were prepared from each horizontal row and vertical column of the matrix (36 total pools) (note: the last horizontal row had only 14 clones). DNA samples were prepared from each pooled culture then used to transfect COS cells. The supernatants from these transfections were assayed using the KG-1 colony assay. Two positives were obtained from this set of transfections: one in a vertical column, the other a horizontal row. The culture common to these pools contained the CSF clone.

Twelve individual clones from this culture were isolated and miniprep DNA was prepared from 10 ml cultures in L-Broth as described above. 1.0 ul samples of DNA from these preparations

were digested with EcoRI and the resulting DNA fragments analyzed by agarose gel electrophoresis. Nine of the twelve clones had a common approximately 750 base pair insert. The DNAs from four of these clones and the remaining three clones were introduced into M6 COS cells as described above. The supernatants from these transfections were assayed using the KG-1 assay as well as the bone marrow assay for CSF. The four clones which each contained the 750 base pair fragment all directed the expression by the M6 COS cells of high levels of CSF activity as detected in either assay while the other three clones did not. Thus, the coding region for CSF must be located within the 750 base pair insert.

The DNA sequence coding for CSF was removed from the transformation vector in the positive clone by digestion with EcoRI and sequenced using standard dideoxy sequencing methods after subcloning fragments into M13 vectors to obtain the sequence illustrated in Fig. 1. The plasmid, p91023(B) - CSF, that was first shown to direct CSF expression in COS cells has been designated pCSF-1. This plasmid has been deposited with the American Type Culture Collection in a strain of E. coli - MC1061 under the deposit number ATCC 39754 on July 2, 1984.

#### Step 7. Expression of CSF protein

M6 COS monkey cells transformed with vector p91023(B) containing CSF/cDNA as isolated in Step 6 are grown as described in Step 6 to produce CSF protein in the culture medium. Namely, one mg of this DNA (pCSF-1) was dissolved in 1 ml of

0.1 M Tris, pH 7.3 and added to 600 ml of DME containing 2 mM glutamine, 100 U/ml streptomycin, 100 ug/ml penicillin (P/S) and 0.25 mg/ml DEAE Dextran (Molecular weight 500,000 from Pharmacia). The 600 ml of DNA DEAE Dextran solution is added to the M6 COS cells in the cell factory and incubated at 37° for 12 hours. After the incubation, the cells are rinsed once with 900 ml of SF DME then incubated for 2.5 hours with 600 ml of DME containing 0.1 mM chloroquin, 10% HIFCS, 2 mM glutamine and P/S. After aspirating the chloroquin containing medium, the cells are rinsed with SF DME and fed 1500 ml of DME with 10% HIFCS. After 30 hours the cells are washed with SF DME, the medium is replaced with 800 ml of SF DME and the transfected cells are allowed to condition the medium for 24 hours at 37°C. The conditioned medium is aspirated and replaced with another 800 ml of SF DME. The cells are allowed to condition this medium for 24 hours then the conditioned medium is collected. As soon as possible after harvesting, the conditioned media sample are concentrated 20 fold by pressurized ultrafiltration using the Amicon 2.5 liter chamber with the YM5 membrane (5,000 MW cutoff).

Step 8. Purification of Recombinant CGP

Two hundred ml of concentrated conditioned medium (from 4 liters of starting material - Step 7) was brought to 30% saturation of ammonium sulfate by addition of solid ammonium sulfate and the precipitated protein was removed by centrifugation. The supernatant was brought to 80% saturation of

ammonium sulfate by adding more solid ammonium sulfate and the precipitated protein collected by centrifugation. The pellet was resuspended in 5 ml of 20 mM sodium citrate, pH 6.1, containing 1 M NaCl. The dissolved protein was applied to a 1.6 x 100 cm column of Ultrogel AcA54 equilibrated in the same buffer. The CSF activity eluted from the column with an apparent molecular weight of 19 k Daltons or after about 90 ml. It has been observed that if the gel filtration is performed at low ionic strength, CSF activity is eluted from the column in two positions with apparent molecular weights of about 19 k Daltons and 38 k Daltons, suggesting that GM-CSF may readily form dimers.) The active fractions were pooled and brought to 0.15% TFA (by addition of 10% TFA) and applied to a Vydac C4 column (0.46 x 25 cm) equilibrated in 0.1% TFA. The column was developed with a linear gradient of 0-90% acetonitrile (1 ml/min., 340 ml total) in 0.1% TFA. The CSF activity eluted between 39 and 43% acetonitrile (Fractions 16-20). A 20 ul sample of Fraction 19 was analyzed by SDS polyacrylamide gel electrophoresis (13.5% gel as described by Lammli, Nature 227, 680 (1970)). A single broad protein band with an apparent MW of 18-26 k Daltons was observed. The rather broad size range for CSF is a common feature of glycoproteins and is thought to reflect extensive but variable addition of carbohydrate. Protein from Fraction 19 was submitted to Edman Degradation using the Applied Biosystems gas phase microsequenator. From approximately 20 ug of protein applied, the sequence of the first 16 amino acids was obtained

(A-P-A-R-S-P-S-P-S-T-Q-P-W-E-H). The high yield of this single protein sequence strongly suggested that the CSF protein in Fraction 19 had been purified to homogeneity. Bioassay indicated that Fraction 19 had  $3 \times 10^7$  units per A<sub>280</sub> absorbance units. Since typical proteins in aqueous solution exhibit a range of extinction coefficients of 0.8 to 1.2 A<sub>280</sub> absorbance units per milligram of protein, the specific activity of the purified CSF is between about  $1 \times 10^7$  and about  $4 \times 10^7$  units/mg when assayed using the human bone marrow cell assay.

#### EXAMPLE B

#### CLONING GIBBON CSF

##### Step 1. Preparation of mRNA from Gibbon T-Cells.

A sample of the gibbon T-Cell line designated UCD-MLA 144 was cultured for several weeks in RPMI 1640 (purchased from Gibco) and 20% fetal calf serum (FCS) until there was obtained  $1 \times 10^9$  total cells. The cells were induced to produce high levels of CSF by activation for 24 hours in the presence of 10 nanograms per ml 12-O-tetradecanoyl phorbol 13-acetate (TPA) in RPMI 1640 plus 1% FCS. The cells were harvested by centrifugation (1000 rpm., 5 min.), washed once with phosphate buffered saline (PBS) and finally collected by centrifugation.

Membrane bound polysome (MBP) mRNA was prepared from these cells using the same procedure as described in Example A for the preparation of Mo cell RNA.

Step 2. First Strand cDNA Reaction

6 ug of MBP mRNA (from Step 1) was diluted into a 50 ul cDNA synthesis reaction mixture (see Example A - Step 4) and the reaction initiated by the addition of reverse transcriptase. After incubation for 30 minutes at 42°C, the reaction was stopped by addition of EDTA to 50 mM, and diluted with H<sub>2</sub>O to 100 ul. The mixture was extracted with phenol/chloroform and further extracted with chloroform. The cDNA/RNA hybrids were separated from unincorporated triphosphates by chromatography on a 2 ml Sepharose CL-4B column. The excluded fractions were pooled and the hybrids collected by ethanol precipitations. The final yield was 570 ng.

Step 3. Second Strand cDNA Reaction

The first strand cDNA pellet (Step 2) was resuspended in 50 ml of H<sub>2</sub>O, and second strand synthesis carried out in a standard reaction mixture with *E. coli* Polymerase I, *E. coli* ligase, and RNase H. The reaction was incubated overnight at 16°C and then incubated for 1 hour at 37°C. The reaction was stopped by addition of EDTA and extracted with phenol/chloroform. The cDNA was separated from unincorporated triphosphates by chromatography on a Sepharose CL-4B column, the excluded fractions pooled and the cDNA collected by ethanol precipitation.

Step 4. Recombinant cDNA Preparation

The cDNA pellet (Step 3) was resuspended in 75 ul of H<sub>2</sub>O.

Homopolymeric C "tails" were added to the ends of the cDNA by adding 10 ul of the cDNA solution to a 25 ul standard reaction mixture with terminal transferase, and incubating at 30°C for 5 minutes. The reaction was stopped by the addition of EDTA to 40 mM and heat inactivation at 68°C for 10 minutes. 10 ng of this tailed cDNA was annealed with 50 ng of G-tailed pBR322 (purchased from NEN) in 10 ul of 10 mM Tris, pH 7.5, 1 mM EDTA, and 100 mM NaCl. The annealing reaction was incubated for 10 minutes at 68°C and then for 2 hours at 57°.

#### Step 5. Bacterial Transformation

E. coli strain MC1061, was grown in L-broth, chilled on ice, harvested by centrifugation, and treated with  $\text{CaCl}_2$  to prepare them for transformation. 5 ul of the cDNA annealing reaction was then incubated with 200 ul of the  $\text{CaCl}_2$ -treated bacteria. Fifteen such transformations were performed, using all of the annealed cDNA, and spread on 15 cm, 1% agar L-broth plates containing 10 ug/ml tetracycline. Approximately 1000 colonies grew on each plate.

#### Step 6. Replica Plating

10,000 colonies from the transformation were each picked with a toothpick, transferred to fresh plates (500 per plate in a grid), and grown overnight at 37°C. The colonies were then lifted from each plate by pressing a dry nitrocellulose filter firmly over the surface of the plate. Two replica filters were

prepared from each of these master filters. The master filters were stored at 4°C, and the replica filters treated with base, and baked to prepare them for hybridization.

Step 7. Preparation of  $^{32}P$  Labelled Hybridization Probes

The cDNA insert from pCSF-1 was isolated by digestion with the restriction enzyme EcoRI, and electrophoresis in an agarose gel with Tris acetate and ethidium bromide. The band containing the cDNA fragment was cut from the gel and purified by the glass powder technique.

300 ng of the cDNA fragment was then added to 1 ul of 10 x T4 DNA Polymerase Buffer (0.33 M Tris Acetate, pH 7.9, 0.66 M potassium acetate, 0.1 M Magnesium acetate and 10 mM dithiothreitol), and 3 units of T4 DNA Polymerase (New England Biolabs), and diluted with water to 10 ul. After incubation for 5-10 minutes at 37°C, this mixture was combined with 1 ul 10 x T4 DNA Polymerase Buffer; 1 ul of a 2 mM solution of each of dCTP, dTTP, dGTP; 10 ul of  $^{32}P$ dATP (10uCi/ul, 3,000 Ci/mmol); and 3 units of T4 DNA Polymerase. The reaction was incubated for 20 minutes at 37°C. Then 1 ul of 2 mM dATP was added and the reaction incubated for an additional 10 minutes at 37°C.

The unincorporated triphosphates were separated from the labelled cDNA by chromatography on a Sephadex G100 column. A second probe was prepared from a synthetic oligonucleotide having the sequence:

ATC TGG CTG CAC AG

ERSATZBLATT

which is complimentary to the amino terminus of the CSF coding region. This oligonucleotide was labelled with  $^{32}\text{P}$  dATP at its 5' end using a standard polynucleotide kinase reaction.

Step 8. Isolation of CSF cDNA Clones

In a standard hybridization screening procedure, some 45 clones hybridized with the T4 labelled pCSF-1 cDNA. Of these, approximately 20 also hybridized to the labelled oligonucleotide probe. The coding region of one of these has been sequenced, and the sequence data revealed a number of base substitutions, some of which result in amino acid difference in the expressed protein. These differences are illustrated in Figure 1 above the DNA sequence for the human CSF gene cloned in Example A.

EXAMPLE C

CLONING CSF FROM PERIPHERAL BLOOD LYMPHOCYTE mRNA

Step 1. mRNA Preparation from Peripheral Blood Lymphocytes

Peripheral blood lymphocytes were prepared from four plasmapheresis by-products (purchased from the Red Cross) by fractionation on a Ficoll-Hypaque gradient. The light density in RPMI-1540 in the presence of 5% fetal calf serum, 0.17% phytohemmaglutinin, and 10ng/ml phorbol myristate acetate (PMA) at a density of  $2 \times 10^6$  cells/ml (a total of  $6 \times 10^9$  cells were obtained). The cells were harvested by centrifugation (1000 rpm, 5 min.), washed once with phosphate buffered saline (PBS) and finally collected by centrifugation. Cytoplasmic RNA was

prepared by a gentle lysis procedure in which the cells were resuspended in 50 ml cold Triton lysis buffer (140 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 10 mM Tris, pH 8.6, 0.5% Triton X-100) with 10 mM dithiothreitol (DTT) and 50 units/ml RNasin (purchased from Biotez). This lysate was divided into 2 equal parts and each part was layered over a 10 ml cushion of lysis buffer containing 20% sucrose. The cell nuclei were removed by centrifugation in the cold (4°C, 400 rpm for 5 minutes). The upper layer (cytoplasmic extract) was carefully removed and sodium dodecylsulfate (SDS) was added to a final concentration of 1%. This solution was extracted twice with an equal volume of phenol chloroform (1:1 mixture) and the RNA was precipitated by adding 2.5 volumes of cold ethanol. The precipitated RNA was collected by centrifugation (15 min. at 4000 rpm) and resuspended in 0.01 M Tris, pH 7.5, 1 mM EDTA, 0.25 M NaCl (TE buffer plus 0.25 M NaCl) and reprecipitated by addition of 2.5 volumes of cold ethanol. Finally, the RNA was collected by centrifugation and resuspended in 5 ml of H<sub>2</sub>O. The final yield was 7.5 mg.

Messenger RNA was isolated from the total cytoplasmic RNA by selection on oligo dT cellulose. 2.5 mg of total RNA was heated to 65° for five minutes. NaCl was added to 0.5 M and the RNA was allowed to cool to room temperature. This RNA was passed over a one ml column of oligo dT cellulose equilibrated in TE + 0.5 M NaCl (binding buffer). Unbound RNA was removed by washing the column extensively with binding buffer. Bound messenger RNA was eluted with 3 ml of H<sub>2</sub>O and precipitated by addition of 0.2 ml

of 4 M NaCl and 2.5 volumes of cold ethanol. The precipitated mRNA was collected by centrifugation (30 minutes at 25,000 rpm). The final pellet (approximately 100 ug) was resuspended in 50 ul of H<sub>2</sub>O.

Step 2. First Strand cDNA Reaction

20 ug of PBL mRNA was diluted into a 50 ul cDNA synthesis reaction containing 100 mM Tris, pH 8.4, 140 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, 400 uM each of dATP, dGTP, dCTP, and dTTP, 5 ug of oligo-dT (average size 12-18) as primer, 25 uCi of <sup>32</sup>PdCTP (400 uCi/mmol) and 20 units of the ribonuclease inhibitor RNasin. The reaction was initiated by addition of 60 units of reverse transcriptase at 37°C and incubated for 30 minutes at 42°C. The reaction was stopped by addition of EDTA to 40 mM and extracted with an equal volume of H<sub>2</sub>O saturated phenol. The phenol phase was back extracted with 50 ul of TE buffer. The aqueous phases were pooled. The cDNA/RNA hybrids were separated from unincorporated triphosphates by passing the pooled aqueous phase over a 5 ml Sepharose CL-4B column (purchased from Sigma), equilibrated with TE. The fractions that were excluded from the column were pooled, brought to 250 mM NaCl and the nucleic acids precipitated by addition of 2.5 volumes of cold ethanol. The hybrids were collected by centrifugation for 30 minutes at 40,000 rpm. The final pellet (2.5 ug of cDNA) was resuspended in 50 ul of H<sub>2</sub>O.

ERSATZBLATT

### Step 3. Second Strand cDNA Reaction

Second strand cDNA was synthesized by the combined action of the enzymes E. coli DNA Polymerase I, E. coli DNA ligase and E. coli RNase H. The reaction mixture (50 ul) contained 20 mM Tris, pH 8.0, 4 mM MgCl<sub>2</sub>, 1.2 mM EDTA, 25 uM NAD, 100 uM each of dATP, dGTP, dCTP, and dTTP; and 50 uCi <sup>32</sup>PdCTP (3,000 Ci/mmol). The reaction was performed by adding 3 units DNA polymerase I, 0.5 units DNA ligase, and 0.75 units of RNase H and incubating at 16° for 18 hours, then at 37° for 1 hour, and then stopped by adding EDTA to 40 mM and extracted with an equal volume of phenol. The phenol phase was back extracted with 50 ul TE, the aqueous phases pooled, and the cDNA was separated from the unincorporated triphosphates by chromatography on a Sepharose CL-4B column as described above for the first strand. Based on incorporation of <sup>32</sup>P, the first strand cDNA was quantitatively converted to a double-stranded form.

### Step 4. Recombinant cDNA Preparation

Homopolymeric C "tails" were added to the ends of the cDNA by gently heating 400 ng of cDNA in a 50 ul reaction mixture containing 1 mM 2-mercaptoethanol, 1 mM CoCl<sub>2</sub>, and 9 units of terminal deoxynucleotidyl transferase at 30°C for five minutes. The reaction was stopped by the addition of EDTA to 40 mM and heating to 68°C for 10 minutes. 200 ng of this tailed cDNA was annealed with 500 ng of G-tailed pAT153 (purchased from Amersham) in 100 ul of 10 mM Tris, pH 7.5, 1 mM EDTA, and 100 mM NaCl. The

ERSATZBLATT

annealing reaction was performed at 57° for 2 hours after a 5 minute preincubation at 68°C.

#### Step 5. Bacterial Transformation

The cDNA annealing reaction product was used directly to transform the *E. coli* strain MC1061. A fresh colony of bacteria cells was used to inoculate 50 ml of L-broth and grown for several hours until the optical density at 550 nm was 0.25. The cells were chilled on ice and harvested by centrifugation (2000 rpm for 10 min.). The pellet was resuspended in 10 ml of cold 0.1 M  $\text{CaCl}_2$  and allowed to sit on ice for 10 minutes. The cells were collected by centrifugation (2000 rpm for 5 minutes) and resuspended in 2.5 ml of 0.1 M  $\text{CaCl}_2$ . 10  $\mu\text{l}$  of the cDNA annealing reaction was then incubated with 200  $\mu\text{l}$  of  $\text{CaCl}_2$ -treated bacterial for 30 minutes on ice and then for 2 minutes at 37°C, followed by addition of 0.8 ml of L-broth and final incubation for 30 minutes at 37°C.

Twenty of these transformations were performed, utilizing all of the annealed cDNA. Each transformation mixture was spread onto 1% Agar L-broth plates (15 cm diameter) containing 10  $\mu\text{g}/\text{ml}$  tetracycline. From the twenty transformations a total of 20 such plates were spread and incubated overnight at 37°C. On the average approximately 1,500 bacterial colonies grew on each plate for a total of 30,000 colonies.

Step 6. Replica Plating

The original colonies growing on each plate were transferred to 137 mm nitrocellulose filters by pressing a dry filter on top of the colonies and lifting them off the plate. Two identical replicas were prepared from each original filter by standard replica plating methods, in which each original filter was carefully placing methods, in which each original filter was carefully placed colony side up on a sterile square of filter paper (Whatman 3 MM) resting on a square piece of glass. A new pre-wetted nitrocellulose filter was carefully aligned on top of the master filter, covered with a second sterile square of filter paper and the complete sandwich then pressed together firmly with a second piece of glass. The sandwiched filters were numbered and 3 pinholes were punched through them asymmetrically so that they could be exactly aligned again in the future. The replica was then removed from the master and placed colony side up on a new tetracycline-containing L-broth agar plate. A second replica was immediately prepared in identical fashion. Each master filter was returned to a plate and all of the plates were incubated at 37° for several hours until the bacterial colonies had reached approximately 1 mm in diameter. The original master filters were stored at 4°C and the replicas prepared for hybridization as described below.

Step 7. Preparation of Filters for Hybridization

Each replica filter (Step 6) was placed colony side up on

filter paper (Whatman 3 MM) soaked in 0.5 M NaOH, 1.5 M NaCl for seven minutes. The filters were transferred to neutralization filter papers, soaked in 1 M Tris, pH 7.5, 1.5 M NaCl, for 2 minutes and then transferred to a second set of neutralization filters for 5-10 minutes. Finally, the filters were placed on filters soaked in SSC buffer (0.015 M Sodium Citrate, 0.15 M NaCl, pH 7.4) for five minutes, air-dried and baked in vacuo at 80°C for 1-2 hours.

#### Step 8. Isolation of CSF cDNA Clones

Duplicate filters were probed with the radioactively labelled pCSF-1 cDNA insert, prepared as described above in Example B. Some 20 colonies hybridized with the cDNA. Twelve of these were picked from the master filter and grown overnight in L-broth for further analysis. Restriction enzyme digests (Pst 1) of DNA samples (rapid prep) from these clones indicated that 3 were nearly full length. One of these has been sequenced. The sequence of the CSF coding region of this clone was identical to the corresponding sequence of pCSF-1 (i.e. having a T at position 365-CSF(Ile)).

#### EXAMPLE D

##### Purification of CSF from Mo Cell Line

Mo serum free conditioned medium (40 liters) was incubated at 55°C for 30 minutes to inactivate the HTLV-II virus associated with the cell line. This medium was concentrated by pressurized

ERSATZBLATT

ultrafiltration using the Pellicon Cassette with membrane PTGC (1.5 square feet) which has a 10,000 molecular weight cut-off. The protein was further concentrated by ammonium sulfate precipitation (80% saturation). The final protein pellet (800 mg) was resuspended in 100 ml of 20 mM Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), pH 7.4, and dialyzed against the same buffer (3 times with 4 liter changes each time). The dialyzed protein was applied to a 2.5 x 10 cm column of DEAE (diethylaminoethyl)-ultracel equilibrated in the same buffer. The column was washed with 800 ml of 20 mM Tris-HCl, pH 7.4, then the CSP activity eluted with 800 ml of 20 mM Tris-HCl, pH 7.4, containing 0.12 M NaCl. 10 ml fractions were collected and assayed for CSP. The active fractions (3) were pooled, and concentrated 6 fold (to 5 ml) by pressurized ultrafiltration (Amicon YM5 membrane, 5,000 molecular weight cut-off). The concentrated sample from the DEAE column was applied to a 1.6 x 100 cm ACA44 ultracel (an acrylamide agarose ultracel having 10 to 130 k Dalton fractionation) column equilibrated in 20 mM N-2-hydroxyethylpipocazine-N-2-ethane sulfonic acid (HEPES), pH 7.4, 50 mM NaCl, and 0.01% polyethylene glycol (PEG-8000). CSP activity eluted from the column with an apparent molecular weight of 30 k Daltons. The active fractions were pooled and brought to 0.15% (v/v) trifluoroacetic acid (TFA) by addition of 10% TFA and applied to a Vydac C<sub>4</sub> reverse phase column (1 x 25 cm). The column was developed with a linear gradient of 0-90% acetonitrile in 0.1% TFA (v/v) at 4 ml/min (1,000 ml total). The CSP activity eluted

at approximately 47% (v/v) acetonitrile. The pooled active fractions were brought to 0.05% (v/v) heptafluorobutyric acid (HFBA) by addition of one half volume of 0.15% (v/v) HFBA and applied to a Vydac C4 column (0.46 x 25 cm) equilibrated in 0.15% (v/v) HFBA. The column was developed with a linear gradient of 0-90% (v/v) acetonitrile in 0.15% (v/v) HFBA at 1 ml/min. (340 ml total). The CSF activity eluted at about 53% (v/v) acetonitrile. Fractions 37-44 (1 ml each) were found to be active. 0.15 ml of fraction 40 was concentrated 4 fold (using the SAVANT Speed Vac Concentrator) and 40  $\mu$ l of 2 x SDS gel sample buffer added (0.125 M Tris-HCl, pH 6.8, 4% SDS, 20% glycerol and 0.004% Bromophenol blue). These samples were boiled for 2 minutes and applied to a 13.5% Lammli, U. Nature 227, 680 (1970) SDS gel (See Figure 2). Fraction (#40) was determined to have 110,000 bone marrow CSF units/ml. This corresponds to about  $3.0 \times 10^7$  units per A<sub>280</sub> absorbance unit. Since typical proteins have extinction coefficients ranging between 0.8 and 1.2 A<sub>280</sub> unit per milligram, the purified CSF had a specific activity in the range of about  $1 \times 10^7$  to about  $4 \times 10^7$  units per mg in the bone marrow assay. A 1  $\mu$ g sample of purified GM-CSF was submitted to Edman Degradation using the Applied Biosystems Gas Phase Microsequenator. The sequence of residues 3 through 5 was determined to be Ala Arg

Ser.

EXAMPLE E

Cotransformation and Amplification of CSF Sequence in CHO Cells Plasmid p91023(B)-CSF was introduced into CHO DHFR deficient cells DUKX-B11 (Chasin & Urlaub PNAS 77:4216, 1980) by protoplast fusion as described (Sandri-Goldin et al. Mol. Cell. Bio. 1 743-752, 1981). The growth and maintenance of the CHO cells has been described (Kaufman & Sharp, J. Mol. Bio. 150 601-621 1981). For protoplast fusion, p91023(B)-CSF-1 was introduced into E. Coli HB101 and bacteria grown in 50 ml of m9 salts containing 0.5% casainino acids, 0.4% glucose, 0.012% MgSO<sub>4</sub>, 5ug/ml thiamine, and 10 ug/ml tetracycline to an absorbance of 0.6 at 600 nm. Chloramphenical was added to 250 ug/ml and the culture incubated at 37°C for an additional 16 hours in order to amplify the plasmid copy number. The cells were centrifuged at 3,000 x g for 10 min. at 4°C and suspended in 2.5 ml of chilled 20% sucrose in 50 mM Tris-Cl pH8.0. Lysozyme was added (0.5 ml of a 5mg/ml solution in 0.25M Tris-Cl pH8.0) and the mixture held on ice for 5 min. EDTA (1 ml of 0.25 M EDTA pH8.0) was added for an additional 5 min. on ice, and then 1.0ml of 0.05 M Tris-Cl pH8.0 was added slowly. The suspension was incubated for 15 minutes at 37°C until the bacteria were converted to protoplasts. The suspension was then slowly diluted with 20ml of prewarmed medium containing 10% sucrose and 10mm MgCl<sub>2</sub> and held at 37°C for 15 min. The solution of protoplasts (approximately 10<sup>9</sup>/ml) was added to CHO, DHFR deficient DUKX-B11 cells in a 6-well plate (approximately 1X10<sup>6</sup> cells/well) at a ratio of approximately 1-2X10<sup>4</sup> protoplasts/cell and the protoplasts were pelleted onto the cells by centrifuging at 2000 RPM for 8 min. in a swinging microtiter dish rotor of an IEC Model K centrifuge. After centrifugation, the super-

natant was removed by aspiration. A 2 ml amount of polyethylene glycol solution 50 g of PEO - 1450, (Baker Chem. Co.) in 50 ml of medium was added to each well of the 6-well plate). The cells were again centrifuged at 2000 RPM for 90 seconds, the polyethylene glycol solution removed, and the plates rinsed 3 times with 4ml of medium/well. Cells were then trypsinized, suspended in 10ml media containing 10% fetal calf serum, and centrifuged in a conical tube at 500 RPM in a clinical centrifuge. Pelleted cells from 3 wells were pooled and plated into a 10 cm tissue culture dish. Fresh medium containing 100 ug/ml of kanamycin, thymidine, adenoxine, deoxyadenosine, penicillin and streptomycin and 10% dialyzed fetal calf serum was added to each plate. The kanamycin was included to prevent the growth of any bacteria which had escaped conversion to protoplasts.

Two days later the cells were subcultured 1:15 into alpha-media with 10% dialyzed fetal calf serum, penicillin and streptomycin, but lacking the nucleosides. Cells were then fed again with the same selective media (lacking nucleosides) after 4-5 days.

Colonies appeared 10-12 days after subculturing into selective media. Two schemes for methotrexate (MTX) selection and amplification have been followed. In the first scheme, single independent cloned transformants were isolated on the basis of DHFR expression and subsequently each clone was propagated under conditions to amplify the copy number of the foreign DNA i.e., growth in increasing concentrations of methotrexate. In the second scheme a pool of multiple independent transformants was isolated on the basis of DHFR expression and propagated together under conditions to

amplify the foreign DNA, i.e. growth in increasing concentrations of methotrexate. Then individual clones were isolated from the mass selected population and analyzed for GM-CSF expression. Those clones exhibiting highest levels of GM-CSF expression were grown again under conditions to further amplify the foreign DNA (i.e. growth in increasing concentration of methotrexate in the culture media).

In one experiment, seven DHFR<sup>+</sup> transformants were pooled into alpha medium lacking nucleosides. These cells were subsequently grown in stepwise increasing concentrations of MTX starting at 0.02 uM then steps to 0.1, 0.5 and 2.0 uM. MTX. When assayed for GM-CSF activity in the KG-1 cell assay, these cells produced from 3,000 to 12,000 units per ml. The selected population was cloned 0.5 uM-MTX and in 2.0 uM-MTX. Clones obtained in 0.5 uM-MTX (010, D2, and B6) were subsequently selected for growth in 2.0 uM-MTX. When assayed for GM-CSF activity in the KG-1 cell assay, the cloned cell lines produced from 15,000 to 300,000 units per ml of GM-CSF activity. The GM-CSF produced according to this Example has the amino acid sequence given for CSF-Thr in Figure 1.

#### EXAMPLE F

##### EXPRESSION OF GM-CSF IN E. COLI

GM-CSF was expressed in E. coli from vector pTALC-185R, a diagrammatic description of which is provided in figure 6. The GM-CSF encoding sequence begins with the synthetic sequence ATG•CCA•CCA•CCT•CCT•TCT•CCA•TCT•CCA•TCT•ACT, which determines the initial 11 amino acid residues of mature GM-CSF. The remainder of the GM-CSF encoding sequence is

pTALC-185R is identical to that of pCSF-1, nucleotides 97-447, followed by the sequence TAR-TAR-TAG. Immediately following the triple terminator there is the pUC-18 polylinker. The tetracycline resistance gene from pBR322 has been inserted, in the opposite orientation to the CSF gene, 100 bases downstream from the pUC-18 polylinker. The tetracycline resistance gene carries its own promoter. Continuing counterclockwise there is next the gene for  $\beta$ -lactamase followed by the pUC-18 (CoLE1) origin of replication.

The final structural feature of the plasmid before returning to CSF sequences is the PL promoter. This promoter is essentially as described by A. Skatzman and M. Rosenberg (in "Molecular cloning, a laboratory manual" (1982), Cold Spring Harbor Laboratory, page 419). CSF expression is driven by the PL promoter after thermal induction in a suitable E. coli host strain.

The parental strain used for all the strain constructions was W3110 lacI<sup>O</sup>L8 (R. Brent and M. Ptashne PNAS 78 (1981) 4204-4208.

A fragment of  $\lambda$  DNA ( $\lambda$ -nucleotides 34499 to 38214) was integrated into the chromosome of W3110 lacI<sup>O</sup>L8 at the lacZ locus. The integration was performed using an integration vector composed of pBR325 sequences carrying the genes for chloramphenicol and ampicillin resistance as well as the pBR322 replication origin. (F. Bolivar Gene 4 (1978) 121-136). The  $\lambda$  DNA fragment is inserted into the lacZ gene, which itself is present on the plasmid as a fragment extending from the BstEII site in LacI to a TthIII site downstream of lacZ.

Integration of the  $\lambda$  DNA into the chromosomal copy of lacZ was achieved by homologous recombination and lac<sup>+</sup>, ampicillin sensitive, chloramphenicol resistant colonies were found. A second recombinational event leading to the removal of all extra plasmid sequences but leaving the  $\lambda$  DNA fragment integrated was screened for on lactose-MacConkey plates. The initial lac<sup>+</sup>, amp<sup>S</sup>, cam<sup>R</sup> phenotype changed to a lac<sup>-</sup>, amp<sup>S</sup>, cam<sup>S</sup> phenotype following the second recombinational event. The resulting strain was called GL400 and was  $\lambda^R$  at 30° and  $\lambda^S$  at 42°. This phenotype demonstrates the existence of a functional chromosomal copy of the CI<sup>857</sup> allele.

GL400 was rendered lon<sup>-</sup> by PL transduction from a lysate grown on strain SG20252 (lac $\alpha$ 169, ara $\alpha$ 139 rps $\alpha$ 1 lon $\alpha$ 100::Tn10). The Tn10 was cured by screening for Tet<sup>S</sup> on selective media (S. Maloy, W. Nunn J. Bacteriol. 145 (1981) 1110-1112).

The final host strain was called GI413 (lac $\alpha$ <sup>0</sup>L8, LacZ $\alpha$  (CI, REX, N), lon $\alpha$ 100).

pTALC-185R was transformed into GI413. An overnight culture of this strain was grown at 30°C in 5mls of induction medium containing 7ugml<sup>-1</sup> tetracycline. Induction medium contains, per liter:

20g Casainino Acids

6g Na<sub>2</sub>HPO<sub>4</sub> 7H<sub>2</sub>O

3g KH<sub>2</sub>PO<sub>4</sub>

0.5g NaCl

1g NH<sub>4</sub>Cl

1% glycerol  
2mg vitamin B1  
2mg  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$   
0.2g  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$

This medium (25mls), containing  $7\text{ugml}^{-1}$  tetracycline, was innoculated with 125ul of the overnight culture and shaken at  $30^\circ\text{C}$  in a water bath until the culture reached a density of  $A_{550} 0.5$ . It was then rapidly moved to a  $40^\circ$  water bath and shaken for a further 2 hours to allow synthesis of GM-CSF. Cells were harvested and checked for their content of CSF by SDS-polyacrylamide gel electrophoresis. Under these conditions GM-CSF accumulates to approximately 5% of the cellular protein.

#### EXAMPLE G

#### Expression of GM-CSF in *Saccharomyces Cerevisiae*

##### A. Vector Construction

A plasmid was constructed which contained the gene for an enzyme in the uracil biosynthetic pathway (URA3) as a selection gene and the 2u origin of replication. This plasmid was derived from Ylp5 (Botstein et al., Gene 8, pp. 17-24 (1979)) with the addition of a fragment containing the origin of replication from the 2 micron plasmid of yeast.

##### B. Isolation of the gene for Glyceraldehyde Phosphate

### Dehydrogenase (GPDH)

Two genes for GPDH have been isolated from yeast (Holland and Holland Journal of Biological Chemistry 255 pp 2596-2605 (1980)). An oligonucleotide probe synthesized from the published sequence was used to isolate a GPDH gene from a plasmid library of yeast genomic DNA by standard methods. A plasmid containing the entire GAP491 gene has been deposited previously (ATCC No. 39777).

### C. Preparation of the glyceraldehyde phosphate dehydrogenase promoter for heterologous gene expression

A plasmid was constructed which allows for the natural spacing of the GPDH promoter from the start of the desired heterologous structural gene. This was accomplished by introducing a KpnI site immediately adjacent to the initiator methionine codon of the GPDH structural gene. The promoter "cassette" was then inserted into the yeast expression vector YOp1.

### D. Isolation of the gene for $\alpha$ factor

A gene for the  $\alpha$  factor mating pheromone has been isolated from yeast (Kurjan and Herskowitz Cell, Vol. 30, pp. 933-943 (1982)). An oligonucleotide probe synthesized from this sequence was used to isolate the gene from a plasmid library of yeast genomic DNA by standard methods.

### E. Preparation of the CSF Expression Plasmid

ERSATZBLATT

From the elements described above, and the human CSF gene, an expression vector (AJ14, Figure 7) was constructed by standard methods. In this vector, the natural leader sequence of CSF has been removed and the sequence coding for mature CSF has been inserted adjacent to the  $\alpha$  factor pre-pro sequence. The junctions between the GPDH promoter,  $\alpha$  factor pre-pro sequence, and mature CSF sequence are precisely (below) and have been confirmed by dideoxynucleotide sequencing. AAATAAACAAAATG.CGTTTCCTTCA..... AAA AGA GAG GCG GAA GCT.GCA CCC GCC CGC TCG...

F. Expression of GM-CSF

The plasmid AJ14 was transformed into a strain of *Saccharomyces Cerevisiae*. Cells were cultured to produce CSF.

In accordance with the above the following specific embodiments of the invention are i.a. envisaged and form part of the invention:

1. A method for preparing and isolating a transformation vector containing CSF/cDNA, said method comprising:
  - preparing RNA from a cell that produces CSF;
  - preparing polyadenylated messenger RNA from said RNA;
  - preparing single stranded cDNA from said messenger RNA;
  - converting the single stranded cDNA to double stranded cDNA;
  - inserting the double stranded cDNA into transformation vectors and transforming bacteria with said vector to form colonies;
  - picking pools of 200 to 500 colonies each and isolating plasmid DNA from each pool;
  - transfecting the plasmid DNA into suitable host cells for expressing CSF protein;
  - culturing the transfected cells and assaying the supernatant for CSF activity; and
  - selecting CSF positive pools and screening the colonies used to make the pool to identify a colony having CSF activity.
2. The method of 1 wherein said cell that produces CSF is a T lymphocyte cell.
3. An eukaryotic cell that secretes protein, said cell being transformed by the expression vector prepared by the method of 1.

ERSATZBLATT

4. The cell of 5 wherein said cell is a mammalian cell.

5. The method of 1, further comprising isolating the DNA coding for CSF protein from the colony having CSF activity, and inserting said DNA coding for CSF into an expression vector having a promoter heterologous to the CSF DNA.

6. A cell transformed by said expression vector prepared by the method of 5.

7. The cell of 6 wherein said cell is prokaryotic.

8. The cell of 6 wherein said cell is eukaryotic.

9. cDNA that codes CSF.

10. cDNA having the nucleotide sequence illustrated in Fig. 1.

11. An expression vector comprising cDNA coding for CSF.

12. An expression vector comprising a nucleotide sequence illustrated in Fig. 1.

13. A transformed cell comprising the expression vector of 11 or an allelic variation thereof.
14. A transformed cell comprising the expression vector of 12.
15. The cell of 13 wherein said cell is prokaryotic.
16. The cell of 14 wherein said cell is prokaryotic.
17. The cell of 13 wherein said cell is eukaryotic.
18. The cell of 14 wherein said cell is eukaryotic.
19. The cell of 17 wherein said cell is a yeast cell.
20. The cell of 18 wherein said cell is a yeast cell.
21. The cell of 17 wherein said cell is a mammalian cell.
22. The cell of 18 wherein said cell is a mammalian cell.
23. The cell of 17 wherein said cell is an insect cell.

24. The cell of 18 wherein said cell is an insect cell.

25. CSF protein made by expressing cDNA coding for CSF in a transformed cell.

26. The protein of 25 wherein said cell is a prokaryotic cell.

27. The protein of 25 wherein said cell is a eukaryotic cell.

28. The protein of 27 wherein said cell is a yeast cell.

29. The protein of 27 wherein said cell is a mammalian cell.

30. The protein of 27 wherein said cell is an insect cell.

31. CSF protein substantially free of the protein of native origin.

32. The CSF protein of 31 substantially free of protein of human origin.

33. CSF protein substantially free of glycosylation.

34. CSF protein glycosylated by expression of cDNA coding for CSF protein in a transformed eukaryotic cell.

35. The CSF protein of 34 wherein said eukaryotic cell is a mammalian cell or an insect cell.

36. CSF protein made by expressing cDNA having the nucleotide sequence illustrated in Fig. 1.

37. The protein of 36 wherein said cell is a prokaryotic cell.

38. The protein of 36 wherein said cell is a eukaryotic cell.

39. The protein of 38 wherein said cell is a yeast cell.

40. The protein of 38 wherein said cell is a mammalian cell.

41. The protein of 38 wherein said cell is an insect cell.

42. Human CSF protein having substantially the amino acid sequence illustrated in Fig. 1 or an allele variation thereof.

43. The CSF protein of 42 wherein said CSF is CSF(Thr).

44. The CSF protein of 42 wherein said CSF is CSF(Ile).

45. The CSF protein of 42 wherein said CSF is Met-CSF.

46. The CSF protein of 42 wherein said CSF is a mixture of CSF and Met-CSF.

47. A therapeutic composition for the treatment of myelo-suppression comprising a myelo-suppression treatment amount of CSF protein in a pharmacological carrier.

48. A method for the treatment of mammals having myelo-suppression, said method comprising treating said mammal with CSF protein.

49. The method of 48 wherein said treating step comprises intravenously injecting said mammal with a therapeutic composition comprising CSF protein in a pharmacological carrier.

50. A method of treating mammals to increase the number of circulating granulocytes, said method comprising treating said mammal with an effective amount of CSF protein.

51. Gibbon CSF having substantially the amino acid sequence illustrated in Fig. 1 or an allele variation thereof.

52. The CSF protein of 51 wherein said CSF is CSF(Thr).

53. The CSF protein of 51 wherein said CSF is CSF(Ile).

54. The CSF protein of 51 wherein said CSF is Met-CSF.

55. The CSF protein of 51 wherein said CSF is a mixture of CSF and Met-CSF.

56. A method for purifying CSF protein from a mixture of proteins suspended in aqueous medium, said method comprising:  
precipitating the protein with ammonium sulfate at 80% saturation to form a pellet containing the CSF protein;  
resuspending the pellet in a buffered solution at a pH in the range of about 6 to about 8;  
applying the buffered solution containing CSF to a chromatographic column, eluting the CSF activity with the buffered solution containing sodium chloride and collecting the fractions having CSF activity; and

pooling the active fractions, applying the pooled fractions to a C4 reverse phase column and eluting the CSF activity with a 0 to 30% acetonitrile gradient to collect the fractions containing CSF activity.

57. The method of 56 wherein said buffer is selected from tris(hydroxymethyl)aminomethane hydrochloride, N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid and sodium citrate.

58. The method of 56 wherein the chromatographic column is loaded with octylsepharose, diethylamino-ethyl-ultrogel or acrylamide-agarose-ultrogel.

59. The method of 56 wherein prior to applying the pooled fractions to the C4 column, the pooled fractions are treated with the acetonitrile gradient in a trifluoroacetic acid solution or a heptafluorobutyric acid solution, respectively.

60. The method of 59 wherein the concentration of trifluoroacetic acid or heptafluorobutyric acid in the eluting solution is 0.10% or 0.15% (v/v), respectively.

61. The method of 56 wherein the aqueous medium containing CSF protein is first treated with ammonium sulfate at 30% saturation to precipitate protein and the supernatant for the remaining steps.

62. The method of 56 wherein after the step of precipitating the protein with ammonium sulfate to form a pellet, the method comprises:  
resuspending the pellet in a solution of Tris-HCl and dialyzing the resulting solution;  
applying the dialyzed solution to a column of DEAE-ultragel;  
eluting the column with a solution of Tris-HCl containing at least 0.1 M NaCl and collecting the fractions containing CSF activity;  
pooling the active fractions and applying the pooled fractions to a column containing AcA44-ultragel equilibrated with a solution of HEPES containing NaCl and polyethylene glycol;

eluting the column with a solution of HEPES with NaCl and polyethylene glycol and collecting the fractions containing CSP activity;

pooling the fractions containing CSP activity and treating the pooled fractions with trifluoroacetic acid;

applying the trifluoroacetic acid treated pool to a C4 reverse phase column, eluting with a 0 to 90% acetonitrile gradient in a trifluoroacetic acid solution and collecting the fractions containing CSP activity;

pooling the fractions containing CSP activity, treating the pooled fractions with heptafluorobutyric acid and applying the treated solution to a second C4 reverse phase column; and

eluting the second C4 reverse phase column with a 0 to 90% acetonitrile gradient in a heptafluorobutyric acid solution to collect the fractions having CSP activity.

63. The method of 56 wherein after the step of precipitating the protein with ammonium sulfate to form a pellet, the method comprises:

resuspending the pellet in a sodium citrate solution containing NaCl and applying the solution to a column containing acrylamide-agarose-ultragel equilibrated in the same buffer;

eluting the column with a solution of sodium citrate and NaCl and collecting the fractions having CSP activity;

pooling the fractions having CSP activity, treating with trifluoroacetic acid and applying the treated solution to a C4

reverse phase column; and

eluting the column with a 0 to 90% acetonitrile gradient in a trifluoroacetic acid solution to collect the fractions having CSF activity.

64. CSF protein having a specific activity of at least about  $1 \times 10^7$  units/mg in the bone marrow assay.

65. The CSF protein of 64 having a molecular weight in the range of about 15,000 to about 26,000 Daltons.

66. The CSF protein of 64 having a specific activity of at least about  $4 \times 10^7$  units per mg of protein in the bone marrow assay.

67. The method of 56 wherein said CSF is purified from a Mo cell conditioned medium.

68. The method of 56 wherein said CSF is purified from a medium obtained by culturing cells transfected with a recombinant vector containing an expressible CSF gene.

69. The method of 56 wherein said CSF is purified from a medium obtained by culturing COS cells transfected with p91023(B)-CSF.

ERSATZBLATT

70. A method of producing a primate colony stimulating factor (CSF) protein, which comprises inserting a gene coding for such protein into an appropriate vector, transforming said vector containing said gene into eukaryotic or prokaryotic host cells and expressing and isolating said CSF protein.

71. A method according to 70, in which the CSF protein is human granulocyte-macrophage CSF.

72. A method according to 70, in which the CSF protein has the amino acid sequence shown for CSF-thr in Fig. 1.

73. A method according to 70, in which the CSF protein has the amino acid sequence shown for CSF-ile in Fig. 1.

74. A method according to 70, in which the CSF protein is Gibbon ape granulocyte-macrophage CSF.

75. A method according to 70, in which the CSF protein has the amino acid sequence shown for CSF-G in Fig. 1.

76. A method according to 70, in which the CSF protein is a protein corresponding in amino acid sequence to a naturally occurring CSF, except that one or more amino acids has been added, substituted or removed without substantially affecting the biological activity of the natural CSF.

77. A method according to 76, in which the CSF protein corresponds in amino acid sequence to a natural CSF protein except that one or more cysteine residues have been replaced by residues of other amino acids.

78. A method according to 77, in which the CSF protein has the amino acid sequence of a natural CSF except that it is proceeded by a methionine residue.

79. CSF protein of 42, 43, 45, 46 in mature form.

80. CSF protein of 42, 43, 45, 46 containing a signal potentiater start region.

81. CSF protein having 127 amino acids.

82. CSF protein obtainable by expressing E. coli MC1061 deposited in the ATCC under the number ATCC 39754 or the 127 amino acid CSF coding nucleotide sequence of plasmid p91023(B)-CSF deposited therein.

WE CLAIM

1. Recombinant CSF protein.
2. CSF protein having a specific activity of  $1 \times 10^7$  units/mg in the bone marrow assay.
3. CSF protein as claimed in claim 2 which is recombinant CSF protein.
4. CSF protein as claimed in claim 1 which is human granulocyte-macrophage CSF.
5. CSF protein as claimed in claim 1 which has the amino acid sequence shown for CSF-thr in Fig. 1, or CSF-ile in Fig. 1, or CSF-G in Fig. 1.
6. A CSF protein as claimed in [redacted] claim 1 which contains the amino acid sequence as shown in Fig. 1 commencing with Ala-Pro ... or wherein the amino acid sequence commencing Ala-Pro ... is proceeded by a methionine residue.
7. A CSF protein according to claim 1 which is a CSF protein corresponding in amino acid sequence to a naturally occurring CSF, except that one or more amino acids has been added, substituted or removed without substantially affecting the biological activity of the natural CSF.
8. A CSF protein according to claim which is a CSF protein having the amino acid sequence of a natural CSF except that it is proceeded by a methionine residue.

9. A method of producing a CSF protein as claimed in claim 1 which comprises isolating said CSF protein as expressed from eukaryotic or prokaryotic host cells into which has been transformed a vector, said vector having inserted therein a gene coding for said CSF protein.

10. A method as claimed in claim 9 wherein expression occurs from an E. coli, CHO or yeast.

11. A method for preparing and isolating a transformation vector containing CSF/cDNA, said method comprising:

preparing RNA from a cell that produces CSF;  
preparing polyadenylated messenger RNA from said RNA;  
preparing single stranded cDNA from said messenger RNA;  
converting the single stranded cDNA to double stranded cDNA;

inserting the double stranded cDNA into transformation vectors and transforming bacteria with said vector to form colonies;

picking pools of 200 to 500 colonies each and isolating plasmid DNA from each pool;

transfecting the plasmid DNA into suitable host cells for expressing CSF protein;

culturing the transfected cells and assaying the supernatant for CSF activity; and

selecting CSF positive pools and screening the colonies used to make the pool to identify a colony having CSF activity.

12. A method for purifying CSF protein from a mixture of proteins suspended in aqueous medium, said method comprising:

precipitating the protein with ammonium sulfate at 80% saturation to form a pellet containing the CSF protein;

resuspending the pellet in a buffered solution at a pH in the range of about 6 to about 8;

applying the buffered solution containing CSF to a chromatographic column, eluting the CSF activity with the buffered solution containing sodium chloride and collecting the fractions having CSF activity; and

pooling the active fractions, applying the pooled fractions to a C4 reverse phase column and eluting the CSF activity with a 0 to 90% acetonitrile gradient to collect the fractions containing CSF activity.

13. cDNA, or an expression vector coding for CSF according to claim 1.

14. A pharmaceutical composition comprising a CSF according to claim 1, or CSF according to claim 1 for use in therapy.

15. A method of treating infection or granulocytopenia or activating neutrophils in animals which comprises administering CSF of claim 1 to an animal in need of such treatment, or use of CSF according to claim 1 for use in the manufacture of pharmaceutical compositions for such method.

16. A pharmaceutical composition comprising a CSF prepared according to the process of claim 12.

17. A method of treating infection or granulocytopenia or activating neutrophils in animals which comprises administering CSF of claim 12 to an animal in need of such treatment, or use of CSF according to claim 12 for use in the manufacture of pharmaceutical compositions for such method.

18. CSF protein obtainable by expressing *E. coli* MC1061 deposited in the ATCC under the number ATCC 39754 or the 127 amino acid CSF coding nucleotide sequence of plasmid p91023(B)-CSF deposited therein.

20

SCT 66AGB ATB 168 CTG CAB AGC CTG CTG CTC TTG GGC ACT GTB GCC TGC  
 RET Trp Leu Glu Ser Leu Leu Leu Leu Gly Thr Val Ala Cys

80

Ser CSF-G

Arg CSF-G

8

AGC ATC TCT GCA CCC GCC CGC TCB CCC AGC CCC AGC AGC ACB CAB CCC TGG GAG CAT  
 CSF-thr Ser Ile Ser Ala Pro Ala Arg Ser Pro Ser Pro Ser Thr Glu Pro Trp Glu His

140

GTB AAT GCC ATC CAB GAB GCC CGG CGT CTC CTG AAC CTG AGT AGA GAC ACT GCT  
 Val Asn Ala Ile Glu Gln Ala Arg Arg Leu Leu Asn Leu Ser Arg Asp Thr Ala

200

Ile CSF-G

Val CSF-G

8

SCT 6AB ATB AAT GAA ACA GTA GAA GTC ATC TCA GAA ATB TTT GAC CTC CAG GAG  
 Ala Glu RET Asn Glu Thr Val Glu Val Ile Ser Glu RET Phe Asp Leu Gln Glu

260

CCG ACC TGC CTA CAB ACC CGC CTG GAB CTG TAC AAG CAG GGC CTG CAB GGC AGC  
 Pro Thr Cys Leu Glu Thr Arg Leu Glu Leu Tyr Lys Gln Glu Leu Arg Gly Ser

320

CTC ACC AAS CTC AAB GGC CCC TTB ACC ATB ATG GCC AGC CAC TAC AAB CAB CAC  
 Leu Thr Lys Leu Lys Gly Pro Leu Thr RET RET Ala Ser His Tyr Lys Glu Glu His

Ile CSF-Ile

T

TGC CCT CCA ACC CCG GAA ACT TCC TGT GCA ACC CAB ACT ATC ACC TTT GAA AGT  
 Cys Pro Pro Thr Pro Glu Thr Ser Cys Ala Thr Glu Ile Thr Phe Glu Ser

100

380

Thr CSF-G

C

TTC AAA GAB AAC CTG AAB GAC TTT CTG CTT GTC ATC CCC TTT GAC TGC TGB GAB  
 Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Glu

450 460 470 480 490 500

Gly CSF-G

CCA GTC CAB GAG TGA GACCGGCCAG ATGAGGCTGG CCAAGCCGG GAGCTCTCT CTCATGAAAC  
 Pro Val Glu Glu

127

510 520 530 540 550 560 570

A

G

8

AAGAGCTGGA AACTCAGGAT GGTCACTTG CAGGGACCAA GGGGTGGGCC ACATCCATGG TGGBAGTGGC  
 580 590 600 610 620 630 640

T

GGGGACCTGC CCTGGGCCAC ACTGACCTG ATACAGGAT GGCAGAAGAA TGGGAATATT TTAACTGAC  
 650 660 670 680 690 700 710  
 AGAAATCACT AATATTCTATA TATTTATATT TTTAAATAT TTATTTATT ATTATTTAA TTTCATATTC  
 720 730 740 750 760 770 780  
 CATATTTATT CAAGATGTT TACCGTAATA ATTATTATTA AAAATATGCT TCTAAAAAA AAAA

FIG. 2

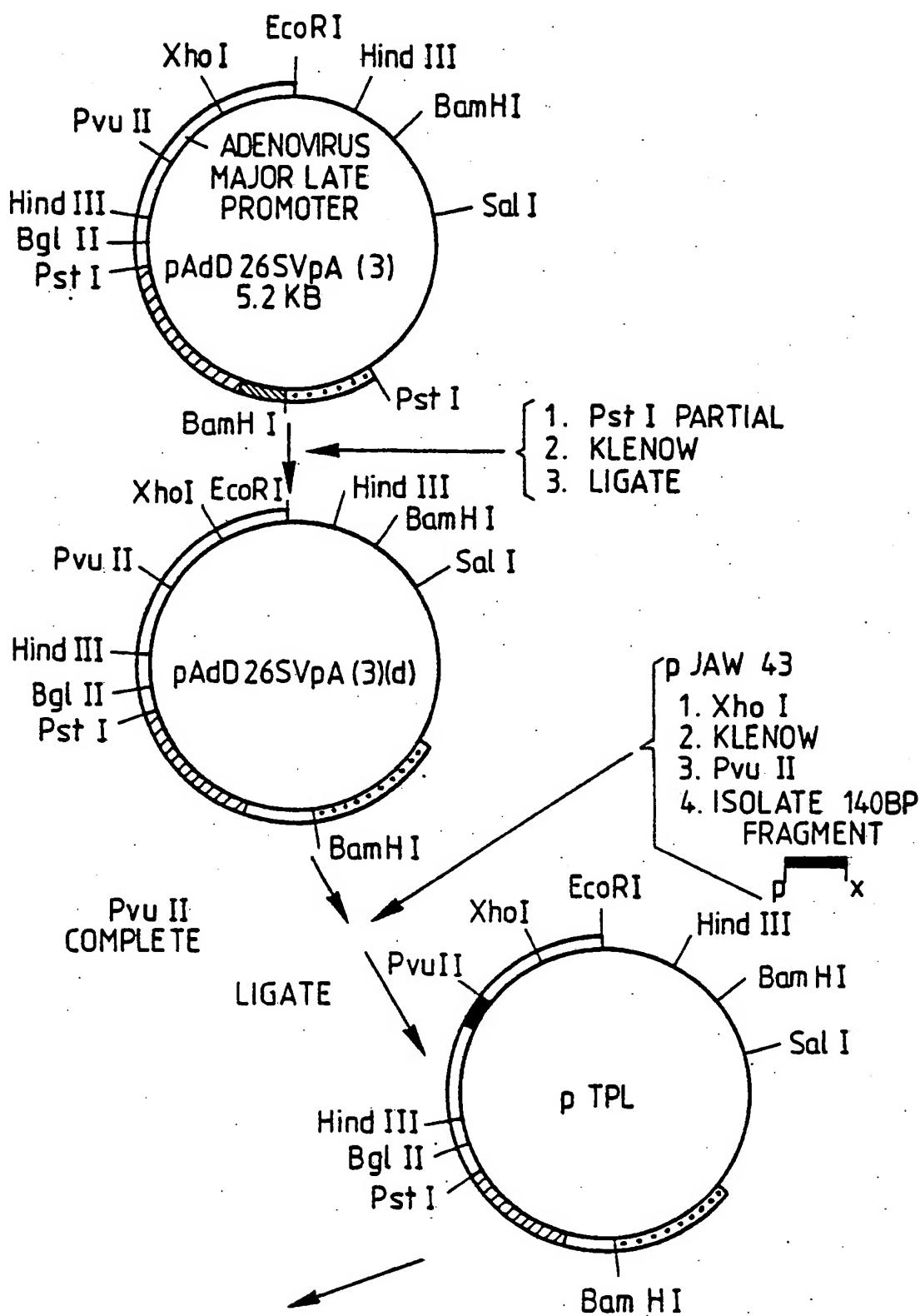
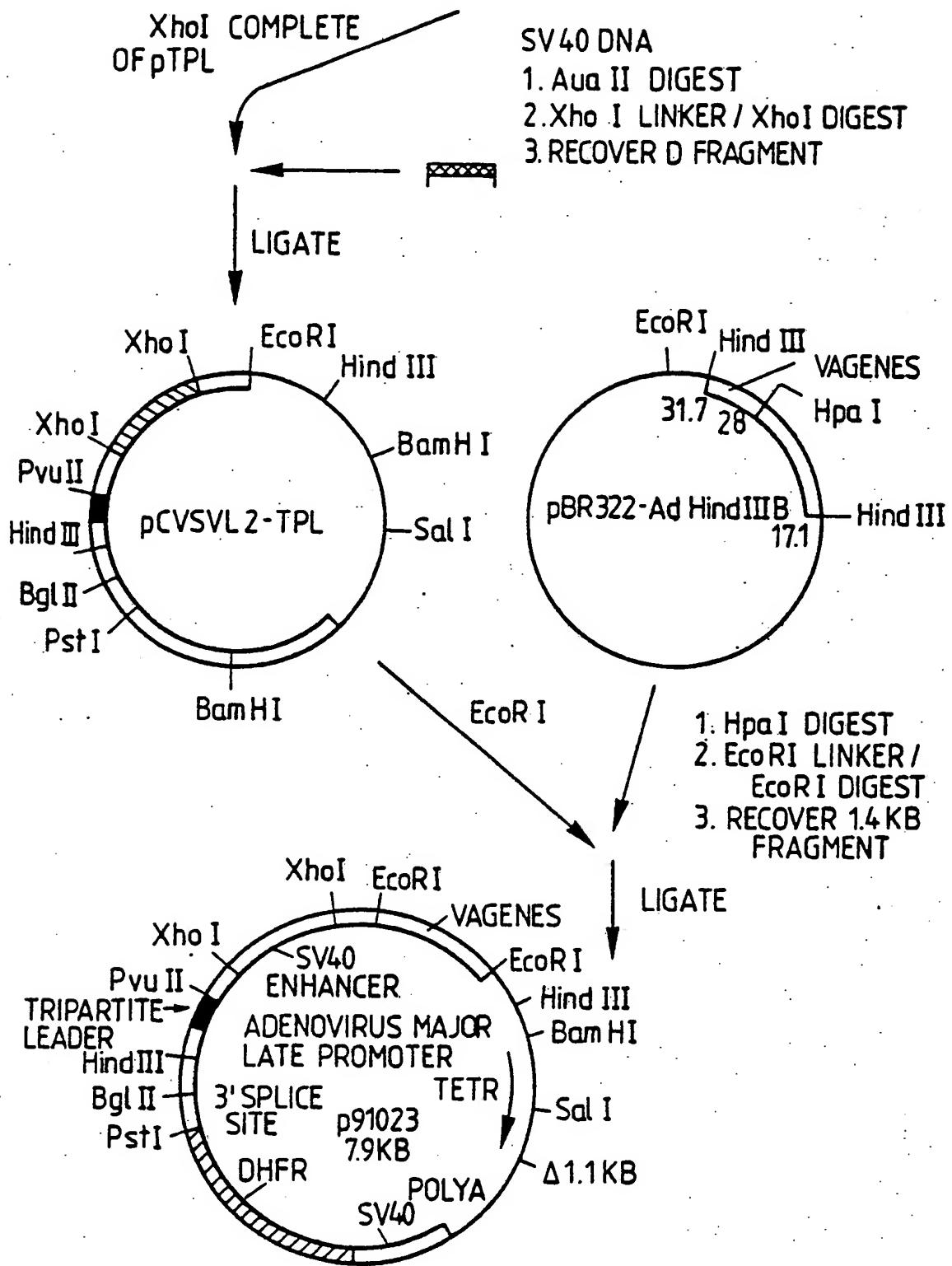
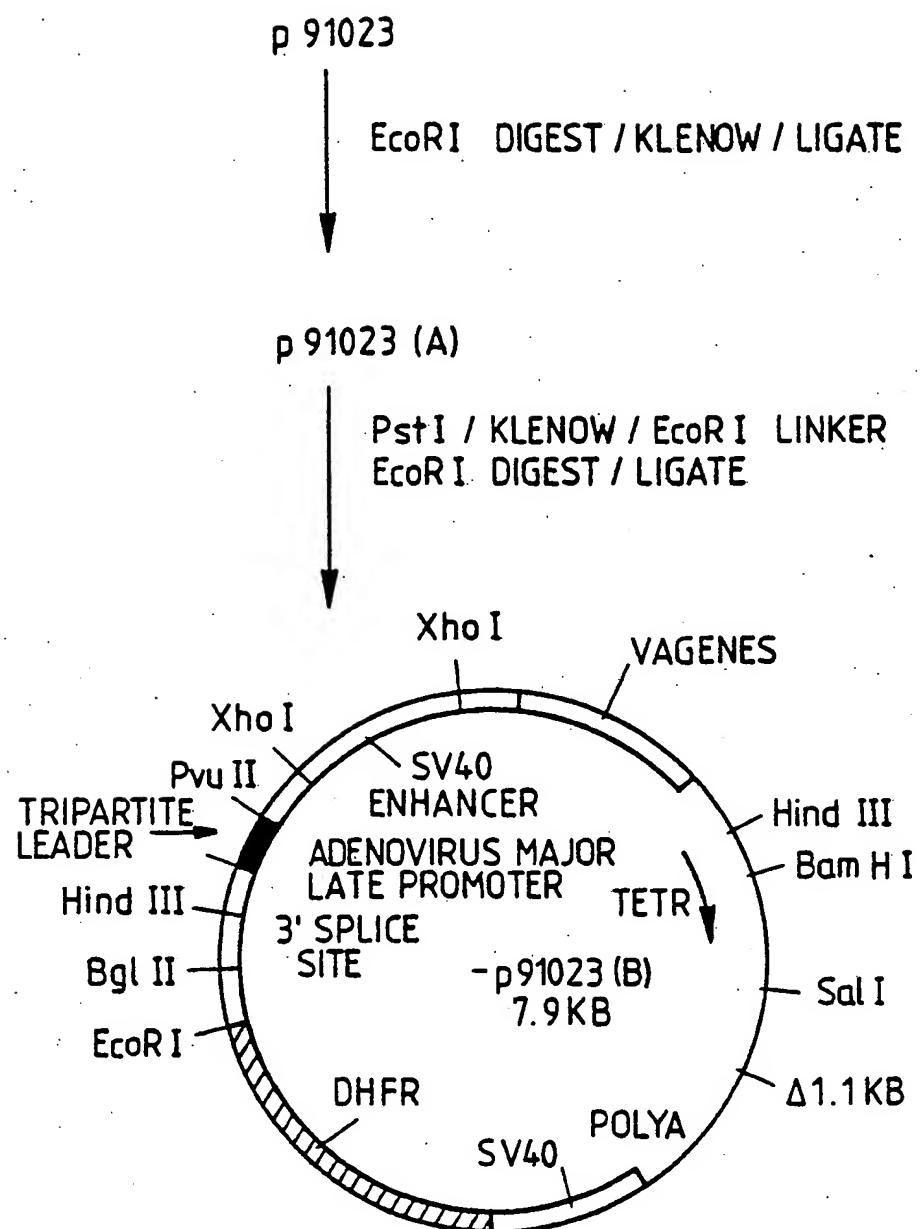


FIG. 3



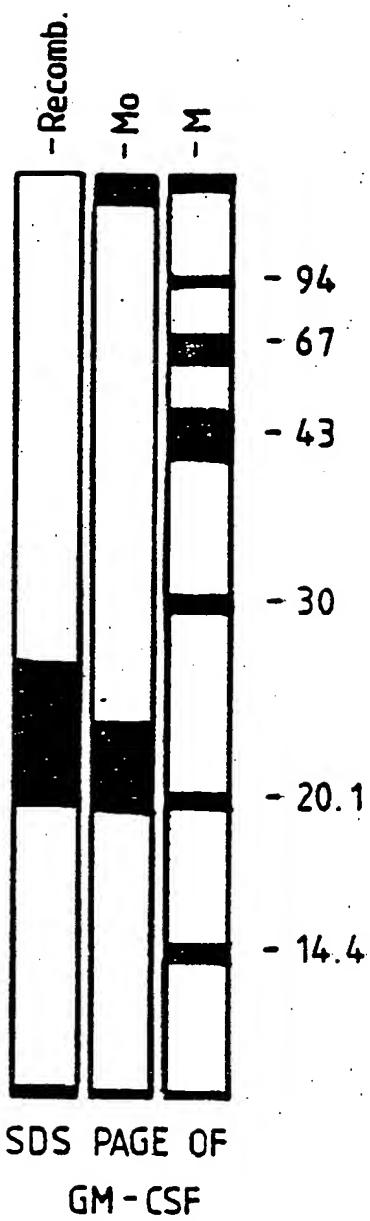
4/7

FIG. 4



5/1

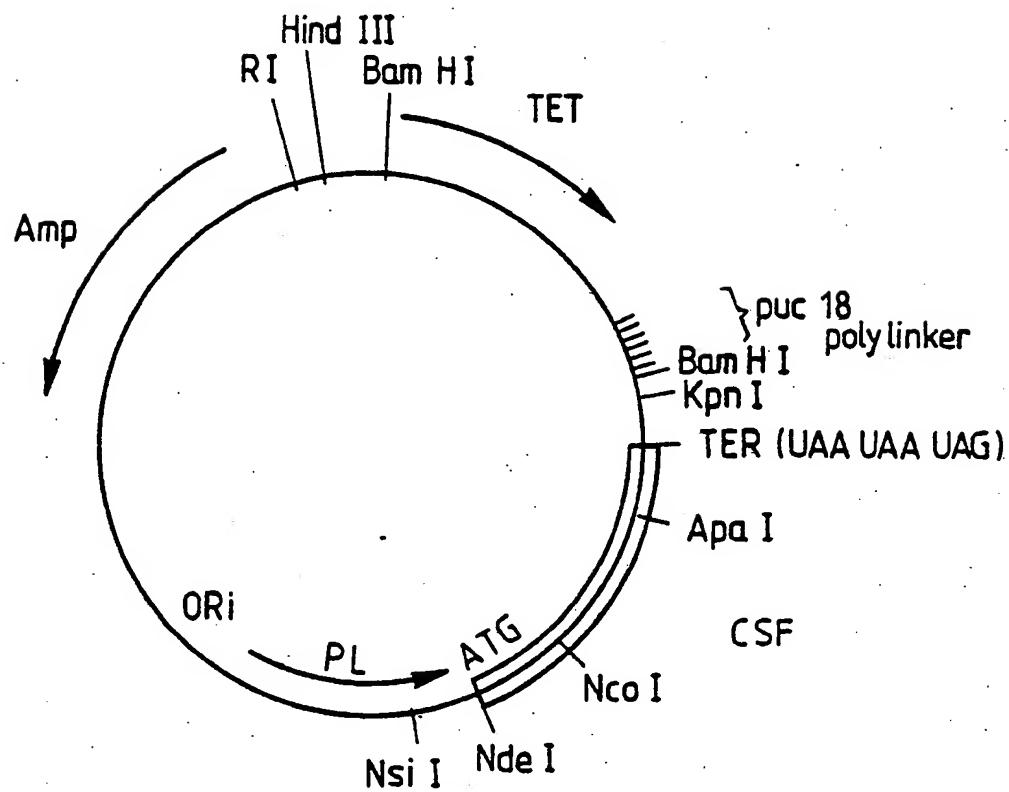
FIG. 5



6/7

FIG. 6

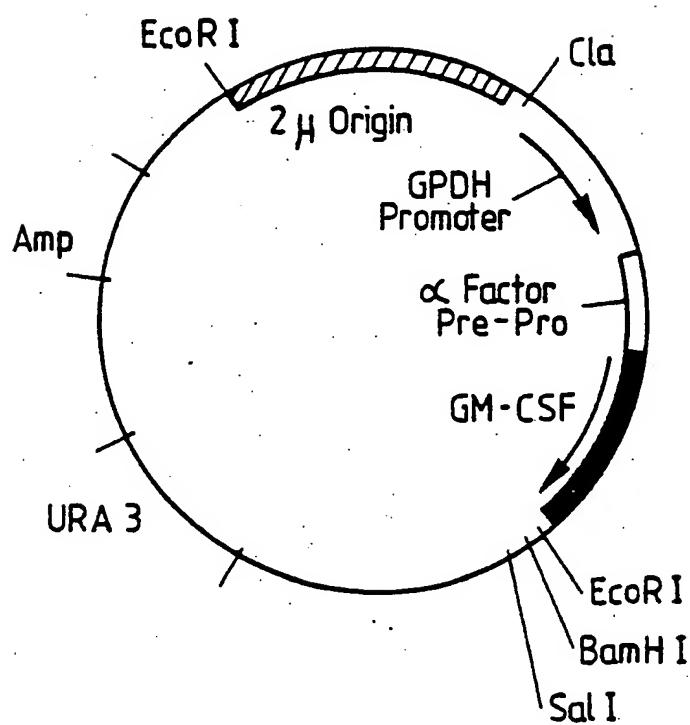
p TALC - 185 R



7/7

FIG. 7

## CSF Expression Plasmid AJ-14



# INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 85/00326

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC<sup>4</sup> : C 12 N 15/00; C 12 P 21/02; A 61 K 37/02

## II. FIELDS SEARCHED

Minimum Documentation Searched<sup>7</sup>

Classification System	Classification Symbols
IPC <sup>4</sup>	C 12 N
	C 12 P
	A 61 K

Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in the Fields Searched<sup>8</sup>

## III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>9</sup>

Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
A	Nature, volume 298, 1 July 1982, Aldons J. Lusis et al.: "Translation of mRNA for human granulocyte - macrophage colony stimulating factor", pages 75-77, see the whole document	1
A	Nature, volume 309, 28 June 1984, Nicholas M. Gough et al.: "Molecular cloning of cDNA encoding a murine haemopoietic growth regulator, granulocyte- macrophage colony stimula- ting factor", pages 763-767, see the whole document	1
A	Nature, volume 307, 19 January 1984, M.C. Fung et al.: "Molecular cloning of cDNA for murine interleukin-3", pages 233-237, see the whole document	1
A	Proc. Natl. Acad. Sci USA, volume 81, Febru- ary 1984, Takashi Yokota et al.: "Isolation and characterization of a mouse cDNA clone that expresses mast-cell growth-factor	

\* Special categories of cited documents:<sup>10</sup>

"A" document defining the general state of the art which is not  
considered to be of particular relevance

"E" earlier document but published on or after the international  
filing date

"L" document which may throw doubts on priority claim(s) or  
which is cited to establish the publication date of another  
citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or  
other means

"P" document published prior to the international filing date but  
later than the priority date claimed

"T" later document published after the international filing date  
or priority date and not in conflict with the application but  
cited to understand the principle or theory underlying the  
invention

"X" document of particular relevance; the claimed invention  
cannot be considered novel or cannot be considered to  
involve an inventive step

"Y" document of particular relevance; the claimed invention  
cannot be considered to involve an inventive step when the  
document is combined with one or more other such docu-  
ments, such combination being obvious to a person skilled  
in the art.

"Z" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search

15th October 1985

Date of Mailing of this International Search Report

14 NOV. 1985

International Searching Authority

EUROPEAN PATENT OFFICE

Signature of Authorized Officer

G.I.M. Spuydenberg

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category*	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	activity in monkey cells", pages 1070-1074, see the whole document --	1
Y	US, A, 4438032 (D.W. GOLDE) 20 March 1984, see claims; column 3-12 --	1,2,12
Y	GB, A, 2092159 (KABUSHIKI KAISHA HAYA-SHIBARA SEIBUTSU KAGAKU KENKYUJO) 11 August 1982, see claims; page 3, lines 31-42 --	1,2,12
Y	EP, A, 0089062 (AJINOMOTO) 9 September 1983, see claims; page 9, line 7 - page 11, line 10 --	1,2,12
P,X	Chemical Abstracts, volume 103, nr. 7, 19 August 1985, Columbus, Ohio, (US) Wong, Gordon G. et al.: "Human GM-CSF: molecular cloning of the complementary DNA and purification of the natural and recombinant proteins", see page 140, abstract number 49105x Science (Washington, D.C. 1883-) 1985, 228(4701), 810-15 (Eng) --	1,4,9,11,13
P,X	Chemical Abstracts, volume 103, nr. 3, 22 July 1985, Columbus, Ohio, (US) Greenberger, Joel S. et al.: "Molecularly cloned and expressed murine T-cell gene product is biologically similar to interleukin-3", see page 452, abstract number 21037c, Exp. Hematol (N.Y.) 1985, 13(4), 249-60 (Eng) --	1
P,X	Proc. Natl. Acad. Sci: USA, volume 82, July 1985 Frank Lee et al.: "Isolation of cDNA for a human granulocyte- macrophage colony-stimulating factor by functional expression in mammalian cells", pages 4360-4364, see the whole document --	1,4,13
P,X	Proc. Natl. Acad. Sci: USA, volume 82, January 1985 Lindsay Sparrow et al.: "Purification and partial amino acid sequence of asialo murine granulocyte-macrophage stimulating factor", pages 292-296, see the whole document -----	1,2,12

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

V.  OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE 1

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.  Claim numbers 15, 17 because they relate to subject matter not required to be searched by this Authority, namely:

SEE RULE PCT 39.1. (iv).

Methods for treatment of the human or animal body by surgery or therapy, as well as diagnostic methods.

2.  Claim numbers ..... because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3.  Claim numbers ..... because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI.  OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING 2

This International Searching Authority found multiple inventions in this international application as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4.  As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

## Remark on Protest

The additional search fees were accompanied by applicant's protest.  
 No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON

INTERNATIONAL APPLICATION NO. PCT/EP 85/00326 (SA 10038)

This Annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 08/11/85

The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A- 4438032	20/03/84	None	
GB-A- 2092159	11/08/82	FR-A;B 2497099 JP-A- 57114525 CH-B- 649786	02/07/82 16/07/82 14/06/85
EP-A- 0089062	21/09/83	JP-A- 58157723	19/09/83